DNA Methylation Analysis of Alzheimer's Disease

Patel, Yogen

Awarding institution:
King's College London

DNA Methylation Analysis of Alzheimer's Disease

Yogen Ghanshyam Patel

Thesis presented for the degree of Doctor of Philosophy

2013

Department of Neuroscience
Institute of Psychiatry
King's College London

ABSTRACT

There is evidence for a role for epigenetic mechanisms in Alzheimer's disease (AD), the most common age-dependent neurodegenerative disorder. The most studied epigenetic mark DNA methylation - the addition of a methyl group to cytosines located in CpG dinucleotides (5mC) - is known to change with aging and may reflect subtle changes in gene expression. Recently a second type of modified cytosine - a hydroxylated and methylated form (5hmC) - has been detected in the brain and maybe linked to the regulation of gene expression. Case-control differences in post-mortem brain DNA methylation were sought by examining both global DNA methylation and DNA methylation of two candidate genes relating to AD risk factors.

Simultaneous assessment of 5mC and 5hmC methylation at a global level indicate hypomethylation of 5mC and hypermethylation of 5hmC in AD brain relative to controls, consistent with the notion that 5hmC serves as an intermediary form for demethylation of 5mC. Age was separately associated with a decrease in LINE1 methylation and an increase in 5hmC methylation.

The comorbidity of depression in AD was explored by assessing the methylation status of the serotonin transporter (SERT) gene promoter across several brain areas and showed tentative associations of disease with SERT CpG methylation. These measurable differences are very small and unlikely to represent any biological plausibility. In a subset of AD patients with additional clinical and behavioural measures there was no effect of SERT 5HTTLPR genotype on DNA methylation.

The hypothesis that amyloid-β deposition in brain is a consequence of amyloid precursor protein (APP) gene over-expression was examined by measuring DNA methylation across the APP gene region. AD status associates with methylation levels of several CpG sites within the 5' region CGI shore and exon 5 of the APP gene. However there are no co-occurring separate associations of total APP protein levels at these CpG sites. This study demonstrates the utility of the Fluidigm microfluidics platform to generate highly parallel bisulphite sequencing/base-pair resolution CpG data.

DECLARATIONS

I declare that this thesis was composed by myself and the research presented is my own.

All the lab work presented in this thesis was conducted by myself, with the exceptions of: (i) the SERT mRNA expression levels used in Chapter 4, which were carried out Dr Petroula Proitsi; (ii) the APP protein levels from Western blotting used in Chapter 5, these were carried out by a M.Sc. student, Jihad Adnan.

All analyses contained in this thesis are my own work, with the exception of the analyses briefly mentioned in Chapter 4, relating to a larger study of clinical and behavioural phenotypes of Alzheimer's disease. All the behavioural measures/variables used in Chapter 4 are also nested within this larger dataset from a study conducted by Dr Petroula Proitsi.

ACKNOWLEDGMENTS

I would firstly like to thank my supervisor Dr John Powell for taking a chance on me and for giving me the opportunity to carry out this research. Without who's guidance and tremendous patience I would have struggled to complete this work. My gratitude also goes to Alzheimer's Research UK for funding this PhD. I would like to thank my second supervisor, Professor Jon Mill and members of his group for all their insightful advice on epigenetic analyses and help in the lab throughout the PhD. For all their help both professionally and personally I wish to thank Dr Michelle Lupton and Dr Petroula Proitsi, your advice and support meant a lot to me. Special thanks go to Dr Kuang Lin for his help with the bioinformatics. Lastly and most importantly, I would like to thank all my family for supporting me, and close friends for their encouragement.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

5hmC	5-hydroxymethylcytosine
5mC	5-methylcytosine
AD	Alzheimer's disease
Aβ	Amyloid-beta
APP	Amyloid precursor protein
BA	Brodmann area
bp	Base pair
C	Cytosine
CGI	CpG island
ChIP	chromatin immunoprecipitation
Chr	Chromosome
CI	Confidence Interval
CpG	Cytosine-guanine dinucleotide
CSDD	Cornell Scale for Depression in Dementia
Ctrl	Control
DMRs	Differentially methylated regions
DNA	Deoxyribonucleic acid
DNMT	DNA methyltransferase
EC	Entorhinal cortex
ELISA	Enzyme-Linked ImmunoSorbent Assay
ELS	Early life stress
ENCODE	The Encyclopedia of DNA Elements
GLUMA	Glucosylation LUminometric Methylation Assay
HATs	Histone acetyltransferases
HDACs	Histone deacetylases
IVS	Intervening sequence
kb	Kilo base
LC-MS	Liquid Chromatography and Mass Spectrometry
LINE1	Interspersed Nucleotide Element-1
LTR	Long Terminal Repeat
MALDI-TOF	Matrix-assisted laser desorption/ionization time-of-flight mass spectrometer
Mb	Mega base
miRNAs	microRNAs
MMSE	Mini Mental Examination Score
mRNA	Messenger RNA
N	Population size
n	Sample size
ncRNAs	Non-coding RNAs
Pb	Lead metal
RNA	Ribonucleic acid
ROI	Region of interest

SERT	Serotonin transporter
SINE	Short INterspersed Elements
SLE	Stressful life events
SNP	Single nucleotide polymorphism
STG	Superior temporal gyrus
TET	Ten-eleven translocation
TF	Transcription factor
TFBS	Transcription factor binding site
TGB	T4 phage β-glucosyltransferase
TSS	Transcription Start Site
UCSC	University of California, Santa Cruz
VC	Visual cortex

ENCODE Transcription Factor ChIP-seq cell types:

1	H1-hESC
A	A549
B	BE2_C
g	GM10847
G	GM12878
g	GM12891
g	GM12892
g	GM15510
g	GM18505
g	GM18526
g	GM18951
g	GM19099
g	GM19193
h	HEK293(b)
H	HeLa-S3
h	HUVEC
J	Jurkat
K	K562
L	HepG2
N	NB4
p	PANC-1
P	PFSK-1
S	SK-N-MC
s	SK-N-SH_RA
U	U87

INTRODUCTION

1.1 AD pathology & genetics

Dementia is characterized by the loss of or decline in memory and other cognitive abilities and as an umbrella term it describes a variety of diseases and conditions that develop as a result of neuronal cell damage or death. Alzheimer's disease (AD) is the most common type of dementia in the elderly and accounts for ~60% of cases (Querfurth & LaFerla 2010). Difficulty in remembering names and recent events is often an early clinical symptom of AD along with apathy. As the disease advances symptoms can include impaired judgment, confusion, disorientation, aggression, depression, trouble with language, and long-term memory loss (Mega et al. 1996; Waldemar et al. 2007). In the later stages, patients have difficulty speaking, swallowing and walking; ultimately leading to the loss of bodily functions and death. The typical duration of disease is 8-10 years, with a range of 1-25 years (Bird 2010), also as the disease progresses the requirement and burden on caregivers can be tremendous. More than 35 million people worldwide and 5.5 million in the United States alone have AD, and these figures are on a steady increase due to a global ageing population, with predictions that 1 in 85 people worldwide will be affected with AD by 2050 (Brookmeyer et al. 2007).

AD was first identified more than 100 years ago (Graeber et al. 1997), but research into its symptoms, causes, risk factors, and treatment has only gained momentum in the past 30 years. Although research has revealed a great deal about AD, the precise physiological changes that trigger the development of AD largely remain unknown (Alzheimer's Association 2012). While many molecular lesions have been detected in AD, the over arching theme to emerge from the data is that an accumulation of misfolded proteins in the aging brain gives rise to the AD phenotype (Querfurth & LaFerla 2010).

The neuropathological hallmarks of AD are the accumulation of amyloid-β (Aβ), which gives rise to extracellular senile plaques and twisted strands of the protein tau, in the formation of intracellular neurofibrillary tangles (NFTs). Together these result in oxidative and inflammatory damage (Wyss-Coray & Mucke 2002; Dyall 2010), which in turn leads to energy failure and synaptic dysfunction (Selkoe 2002; Forero et al. 2006). Overall brains with advanced AD show dramatic shrinkage from neuronal loss and widespread debris from dead and dying neurons.

The major component of NFTs is an abnormally hyperphosphorylated and aggregated form of tau. Normally tau is a soluble protein that promotes assembly and stability of microtubules, whereas hyperphosphorylated tau is insoluble, lacks affinity for microtubules, and self-associates into paired helical filament structures giving rise to NFTs (Alonso et al. 2001). Whereas the inclusions of NFTs in neurons occur in both AD and other neurodegenerative disorders termed tauopathies (Goedert & Jakes 2005), the number of these tangles is a pathologic marker for the severity of AD.

In-life AD diagnosis is based on tests that evaluate behaviour and cognitive abilities, such as the mini-mental state examination (MMSE, Folstein et al. 1975) and often followed by a brain scan if available, however definitive diagnosis rests on neuropathological examination of post-mortem brain after death. Upon post-mortem examination, disease diagnosis and stage is classified using a number of different morphological and histopathology markers.

A widely adopted classification system of disease pathology is the Braak staging system, originally proposed by Braak & Braak (1991). Braak staging uses a six-point scale (I-VI) to primarily reflect the distribution of NFTs within the brain (Figure 1.1). Braak stages I and II are used when NFT involvement is confined mainly to the transentorhinal region of the brain, stages III and IV when there is also involvement of limbic regions such as the hippocampus, and V and VI when there is extensive neocortical involvement (Braak & Braak 1991). While imperfect, the Braak staging system is considered a good anatomical map of AD-related brain dysfunction, and in general agrees with neuropsychological (cognitive decline) and imaging indicators of disease (Giannakopoulos et al. 2003; Whitwell et al. 2007). Whereas, amyloid-β deposition shows far more heterogeneity and is less well correlated with cognitive features in AD (Giannakopoulos et al. 2003; Cupidi et al. 2010). For these reasons Braak staging has remained one of the main disease pathology correlates with which retrospective clinical measures of AD are compared.

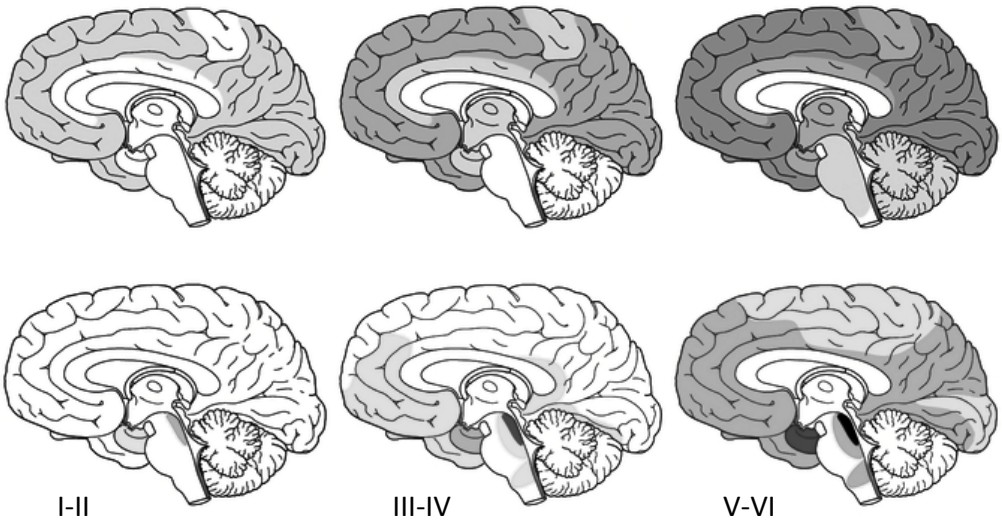


Figure 1.1 Alzheimer's disease progression and distribution of senile plaques & neurofibrillary tangles [amyloid-β plaques (top) and tau neurofibrillary tangles (bottom) spread thorough the brain of AD patients in a specific pattern (left to right). Plaques appear first in the cortical regions and spread to the rest of the brain, whereas NFTs appear in the limbic system (entorhinal cortex, hippocampus, dentrate gyrus) and then spread to the cortical regions. Darker shading indicates protein density over time; the spread of tau NFTs in the bottom pane corresponds to Braak stages I-II, III-IV and V-VI left to right; adapted from Braak & Braak 1991]

The greatest risk factor for AD is advancing age. The age at which people present with AD defines two distinct groups; those with early-onset familial AD (EOFAD, <65) and those with late-onset AD (LOAD, ≥65). Broadly speaking these groups show a genetic dichotomy of cases with Mendelian inheritance versus complex or multifactorial genetic cases and while EOFAD is rare, accounting for ~5% of cases the vast majority of AD cases are late-onset and sporadic in nature.

Currently, autosomal dominant mutations in three genes are known to cause EOFAD, all leading to altered cerebral levels of amyloid-β, the main constituent of the hallmark senile plaques (Hardy 2006). The mutated genes leading to EOFAD are APP (amyloid precursor protein) on chromosome 21 (Goate et al. 1991), PSEN1 (presenilin 1) on chromosome 14 (Sherrington et al. 1995), and PSEN2 (presenilin 2) on chromosome 1 (Levy-Lahad et al. 1995). Aβ is liberated from APP via two enzymatic cleavage events, mediated by β-secretase and γ-secretase (Figure 1.2). The latter is a protein complex consisting of various components, in which the presenilins form the γ-secretase active site. While lifelong over-expression of APP and the resultant overproduction of Aβ in the brains of Down syndrome (trisomy 21) individuals is thought to lead to AD after the age of 40, accounting for <1% of AD cases (Wisiewski et al. 1985; Bird 2010). Though mutations within the tau protein gene, MAPT (microtubule-associated protein tau) on chromosome 17 have been detected in fronto-temporal dementia with Parkinsonism (Rademakers et al. 2004), MAPT mutations do not occur in AD.

The identification of the familial AD genes form the basis of the amyloid cascade hypothesis of AD aetiopathology where an imbalance of Aβ production and Aβ clearance is seen as the initiating event in the disease process, resulting in the aberrant phosphorylation of tau and its intracellular aggregation in the form of NFTs (Hardy & Selkoe 2002).

Despite LOAD having a relatively high heritability of ~70% (Avramopoulos 2009), the only long established unequivocal genetic risk factor has been the ε4-allele of the apolipoprotein E gene (APOE, Saunders et al. 1993; Grupe et al. 2007; Harold et al. 2009; Naj et al. 2011; Genin et al. 2011). The APOE-ε4 allele appears to increase the lifetime risk for AD from 20% to 90% and lowers the age of onset in a dosage dependent manner (Corder et al. 1993; Verghese et al. 2011). The precise mechanism by which APOE exerts this influence remains largely undefined, with indications as a key role in mediating Aβ metabolism and clearance (Mohandas et al. 2009).

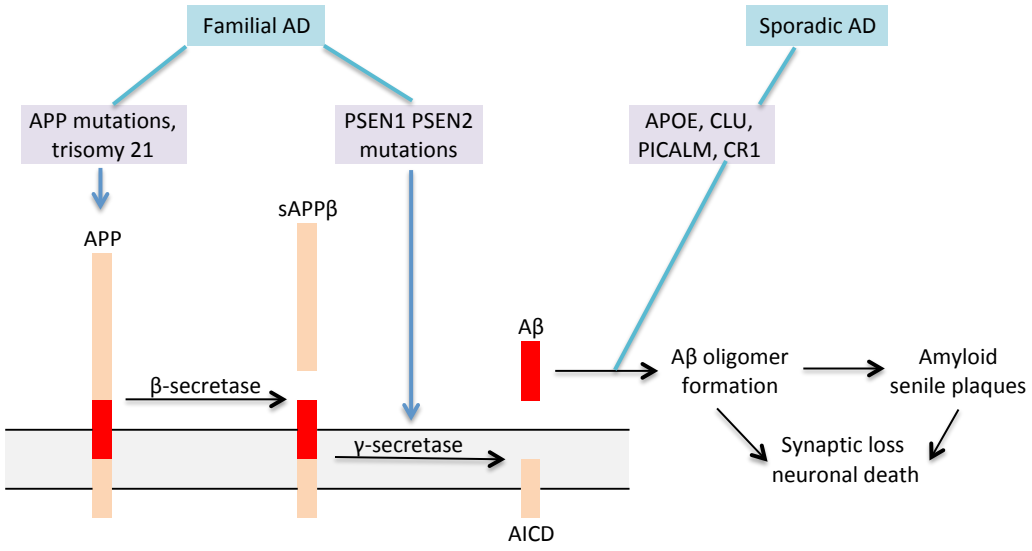


Figure 1.2 The balance between APP processing and clearance in a model of AD pathogenesis [The Aβ peptide is derived via sequential cleavage of the larger precursor, transmembrane protein called amyloid protein precursor (APP). Processing of APP leading to amyloid plaques (amyloidogenic pathway) firstly involves cleavage of APP by β-secretase resulting in the release of a soluble fragment called sAPPβ, then followed by a second cleavage by γ-secretase releasing neurotoxic Aβ into the extracellular space and leaving behind a c-terminal fragment AICD of APP. Oligomer formation of Aβ results in extracellular amyloid plaques giving rise to synaptic loss and neuronal death. Whereby this imbalance of Aβ production and Aβ clearance is seen as the initiating event leading to neurodegeneration and cognitive decline associated with AD. Mutation either in the APP gene or the genes encoding proteins of γ-secretase complex (PSEN1 and PSEN2) account for most of the familial forms of AD, whereas late-onset sporadic forms of the disease are linked to variants in genes (APOE, PICALM, CR1 and CLU) possibly relating to the Aβ metabolism; adapted from van Es & van den Berg (2009)]

Numerous candidate gene, genetic linkage and association studies have been carried out over the past two decades to elucidate the remaining genetic risk to AD (Bertram et al. 2007,

www.alzgene.org). However none of these candidates has been proved to consistently influence disease risk or onset age in more than a handful of samples (Bertram 2009). Reasons for a lack of reproducibility may include: insufficient study power to detect variants with minor contributions; biologic, genetic and allelic heterogeneity; difference in study design and the presence of population substructure (Bettens et al. 2010).

The recent application of genome-wide association studies (GWAS), powered by large international sample collaborations, has led to the identification of many new loci CLU (clusterin), PICALM (phosphatidylinositol binding clathrin assembly protein), CR1 (complement receptor 1), BIN1 (bridging integrator 1), CD33 (sialic acid binding Ig-like lectin 3), CD2AP (CD2-associated protein), ABCA7 (ATP-binding cassette transporter member 7), EPHA1 (ephrin type-A receptor) and MS4A (membrane-spanning 4-domains subfamily A) (Harold et al. 2009; Lambert et al. 2009; Hollingworth et al. 2011; Naj et al. 2011) in addition to APOE, as loci for AD.

CLU and CR1 are implicated in the pathologic AD pathway involving Aβ and its clearance from the brain, whereas the role of PICALM in AD is unclear (Figure 1.2). Furthermore clusterin has striking similarities to APOE, as they have been both shown to cooperate in suppressing Aβ deposition as well as in modifying Aβ clearance at the blood brain barrier (DeMattos et al. 2004; Bell et al. 2006); while CR1 may be involved in Aβ clearance via the complement pathway.

Further pathway analysis of these GWAS data and biomarker studies have implicated inflammatory and cholesterol metabolic pathways as playing an important role in the aetiology of AD (Candore et al. 2010; Jones et al. 2010; Lambert et al. 2010; Fiala & Veerhuis 2010). In short CLU, CR1, CD33 ABCA7 and EPHA1 are linked with immune system functions; CLU, ABCA7 and APOE to cholesterol metabolism; while BIN1, PICALM, CD33 and CD2AP are related to synaptic dysfunction and cell membrane processes (Piaceri et al. 2013).

Most recently exome sequencing has highlighted rare variants in the TREM2 (triggering receptor expressed on myeloid cells 2) gene that cause susceptibility to LOAD, with an odds ratio similar to that of the APOE ε4 allele (Guerreiro et al. 2013). TREM2 plays a role in mediating proinflammatory innate immune signalling by microglia, further supporting the idea that inflammation is integral to the AD disease process (Neumann et al. 2012). Other processes involved in the pathogenesis of AD include calcium homeostasis (LaFerla 2002), oxidative stress and mitochondrial dysfunction (Forero et al. 2006; Sultana & Butterfield 2010; Querfurth & LaFerla 2010).

1.2 Gene expression changes & risk factors in AD

Studies using microarrays show widespread gene expression changes in AD that relate to a number of different pathways. For instance lymphocytes from AD and healthy aged controls show expression difference for genes involved in the regulation of vascular tone and blood pressure, immune response, apoptosis, cell metabolism and transcriptional regulation (Kalman et al. 2005). More specifically altered gene expression in AD brain (Loring et al. 2001; Dunckley et al. 2006; Liang et al. 2008; Miller et al. 2008), further implicates pathways that underlie neurodegenerative processes. These include aberrant cell cycle events, energy metabolism, synaptic plasticity, inflammatory processes and oxidative stress. Although each mechanism could contribute to disease pathogenesis, to what extent they drive the neurodegenerative process is uncertain.

In addition to genetic factors, many environmental factors have been associated with AD risk. These factors include behavioural, dietary and other environmental factors. For instance epidemiological studies suggest that a low education level, history of head trauma, systemic infections and inflammations may each increase the risk of AD (Mayeux 2003; Van Den Heuvel et al. 2007; Perry et al. 2007). While other environmental risk factors that have been related to AD are dietary deficiencies of vitamin B or folate and exposure to metals or pesticides (Kronenberg et al. 2009; Wu et al. 2008; Mastroeni et al. 2009; Hayden et al. 2010); some environmental factors show a protective effect, such as a higher consumption of fruits, vegetables, omega-3 fatty acids, antioxidants, and a low-caloric diet along with moderate physical exercise (Scarmeas et al. 2009).

Furthermore social factors are also associated with risk of AD, people not cohabiting with a partner, widowed or divorced in mid-life, have an increased risk of cognitive impairment or increased later-life risk for AD compared to married or cohabiting people (Hakansson et al. 2009). The exact causal mechanisms underlying these associations with social environmental exposures is not clear, although it is tempting to speculate that psychological stress may be involved (Chouliaras et al. 2010). Interestingly, while there is a high co-morbidity between AD and depression, it has been suggested that brain changes associated with depressive episodes that compromise the brains ability to cope with stress may constitute risk factors for development of AD (Aznar & Knudsen 2011).

Although several AD genes have been identified, especially for the familial form of AD, the genetic variants associated with LOAD do not fully explain the heritability of the disease and account for only ~40% of the genetic risk (Bird 2010). Further still these risk factor genotypes are neither necessary nor sufficient for LOAD development (Bakulski et al. 2012) as demonstrated by AD twin studies revealing incomplete concordance and variable age at onset even among monozygotic pairs (Gatz et al. 1997; Nee et al. 1999). These observations along with the missing heritability and the proposed role of environmental factors emphasise the likely importance of non-genetic factors in LOAD aetiology. The occurrence of phenotypic differences in monozygotic twins over time are thought to arise from epigenetic changes induced by different environments or by stochastic events (Fraga et al. 2005), thus it seems plausible that epigenetic changes resulting in altered gene expression may also be involved in the pathogenesis of LOAD.

1.3 Epigenetics

Superimposed upon the DNA sequence is a layer of heritable "epigenetic" information. This epigenetic information is largely stored as chemical modifications to cytosine bases and the histone proteins that package the genome into chromatin (Figure 1.3). These modifications regulate gene expression but do not change the genetic sequence and are crucial to the normal development and cell differentiation. Thus, through epigenetic modification diverse cellular phenotypes and functions required by the body can be achieved using a single genetic code. These effects are typically achieved through restriction of transcriptional access to genes, leading to their repression or silencing. Conversely, release from constrained epigenetic repression can enhance gene expression. Like the DNA sequence, the epigenetic profile of a cell is passed from mother to daughter cell at mitosis. However, unlike the largely immutable DNA sequence epigenetic modifications are dynamic and can be reversed in even fully differentiate brain cells (Weaver et al. 2005; McCarthy et al. 2009). Epigenetically regulated and dysregulated transcription states can give rise not only to different cell types and developmental stages, but also to favorable and unfavorable outcomes for specific cells within the same organ system. Thus, changes in epigenetic regulation can cause some cells to develop structural, physiologic, and metabolic abnormalities, while other cells of the same type remain normal (Mastroeni et al. 2011).

1.3.1 DNA methylation

DNA methylation is the most studied and well characterised epigenetic mark associated with both transcriptional regulation and genome stability. On the whole DNA methylation functions to repress transcription and has a major role in the processes of X-chromosome inactivation and imprinting where monoallelic expression is controlled by DNA methylation marks (Yen et al. 1994; Edwards & Ferguson-Smith 2007). Other roles are though to include hypermethylation of repetitive elements, protecting the genome from reactivation of endoparasitic sequences and the preservation of genome stability (Bird 2002; Rizwana & Hahn 1999).

DNA methylation consists of the addition of a methyl group ($-CH_3$) to the fifth carbon position of cytosines at CpG dinucleotides (5-methylcytosine). DNA methylation is carried out by a family of DNA methyltransferases (DNMTs), where DNMT1 predominantly maintains on-going methylation, with specificity for hemi-methylated CpGs during DNA replication. DNMT3a and DNMT3b catalyse de novo methylation when DNA methylation patterns are laid down during development (Bernstein et al. 2007). The methyl group that is transferred to cytosine by the DNMTs ultimately derives from 5-methyltetrahydrofolate through its interactions with S-adenosylmethionine (SAM) during one-carbon metabolism (Figure 1.4).

CpG dinucleotides are observed at a five-fold less than expected frequency throughout the human genome, owing to spontaneously deamination of 5-methylcytosine to thymine, leading to transition mutations, where C is replaced by T (Sved & Bird 1990). This is the most common single nucleotide mutation in the human genome and leads to the subsequent loss of methylated CpGs in the germ line over time (Sved & Bird 1990; Saxonov et al. 2006). While CpGs are a rare dinucleotide, their distribution throughout the genome is not uniform. CpGs are enriched in certain regions called CpG islands (CGIs) by definition they have an elevated G+C content of more than 55% and an observed/expected CpG ratio over 0.65. CGIs are approximately 500bp in length and overlap the promoter regions of ~70% of all human genes, which tend to be hypomethylated in normal cells (Illingworth & Bird 2009). The hypomethylated state of CGIs is accompanied by an open chromatin state, facilitating gene expression. Conversely, methylation of CGIs has been related to gene silencing, either by direct inhibition of transcription by precluding the recruitment of transcription factors to their target sites, or through the recruitment of methyl-CpG-binding domain (MBD) proteins, which promote the binding of chromatin remodeling complexes (Siegfried et al. 1999; Jones & Takai 2001; Miranda & Jones 2007).

Recently, transcriptional regulation through methylation sites outside of the promoter region CGI, at sequences ±2 kb form the CGI (called CGI shores) was shown to be important for tissue-specific gene expression (Irizarry et al. 2009). It has also been proposed that DNA methylation within gene bodies may have a potential role in transcriptional activity, by regulating alternative promoters and transcriptional efficiency of RNA polymerase II (Maunakea et al. 2010; Lorincz et al, 2004). With regards to tissue-specific methylation patterns, the brain has been reported to have higher levels of methylation compared to other tissues, mainly in repetitive sequences of the genome (Ehrlich et al. 1982; Schumacher 2011). Furthermore tissue-specific differences extend across different brain regions where methylation differences between cerebellum and cortex far exceed inter-individual DNA methylation variation in any given brain area and in general greater tissue-specific DNA methylation is seen at CGI shores and intragenic CGIs as opposed to promoter CGIs (Ladd-Acosta et al. 2007; Davies et al. 2012).

In addition methylation can also occur at non-CpG dinucleotides within the sequence context of CpH, where H is A, C or T and while this is a less researched phenomena, it has been observed to decrease with differentiation of human stem cells and seems to be enriched in gene bodies (Ramsahoye et al. 2000; Lister et al. 2009; Ziller et al. 2011), others have reported non-CpG methylation in adult brain at the sequence context of -CAC- (Varley et al. 2013). Furthermore recent identification of further cytosine modifications (5-hydroxymethylcytosine, 5-formylcytosine and 5-carboxylcytosine) that appear to be at higher levels in the brain relative to other tissues highlights further complexity of DNA modifications. Also these modifications are thought to relate to active demethylation pathways via oxidation by the ten-eleven translocation (TET) enzymes (Kriaucionis & Heintz 2009; Nabel & Kohli 2011). However, the biological implications of these modifications are still unclear and warrant further investigation. Global methylation levels of both 5-methylcytosine and 5-hydroxycytosine will be discussed and examined further in Chapter 3 of this thesis.

Emerging evidence from rodent studies suggests that DNA hypo- and hypermethylation are dynamic events that can occur within post-mitotic neurons (Levenson et al. 2006; Murgatroyd et al. 2009), supporting the notion that dynamic epigenetic mechanisms may help mediate neuronal and synaptic plasticity. For example, gene-specific hypomethylation of hippocampal neurons after DNMT inhibition blocks long term potentiation (Levenson et al. 2006) and fear conditioning (Miller & Sweatt 2007), whereas Weaver et al. (2005) have suggested that early life events such as maternal care alter adult stress responses through sustained DNA methylation changes in rat hippocampal neurons.

Changes to both the DNA and histones impact gene transcription but do so in different ways. DNA methylation blocks transcription factors from binding to the sequence. Conversely, changes to histones occur on specific residues but act globally to relax or tighten the chromatin surrounding a particular gene region and thereby regulate access of the transcription complex. While the focus of this thesis is on DNA methylation in AD, histone modifications and non-coding RNA as epigenetic mechanisms are briefly covered in the following section.

1.3.2 Histone modifications & non-coding RNA

At the most basic level ~1.5 turns of DNA (147bp) is wrapped around a nucleosome core of a histone octamer consisting of 2 copies each of the histones H2A, H2B, H3, and H4 (Luger et al. 1997) as depicted in Figure 1.3. The N-terminal tails of histone H3 and H4 are highly charged and are tightly associated with DNA as it folds around the nucleosome core. The dominant changes to histones are methylation and acetylation of the protruding tails, but can also include ubiquitylation, phosphorylation, and sumoylation (Berger 2007).

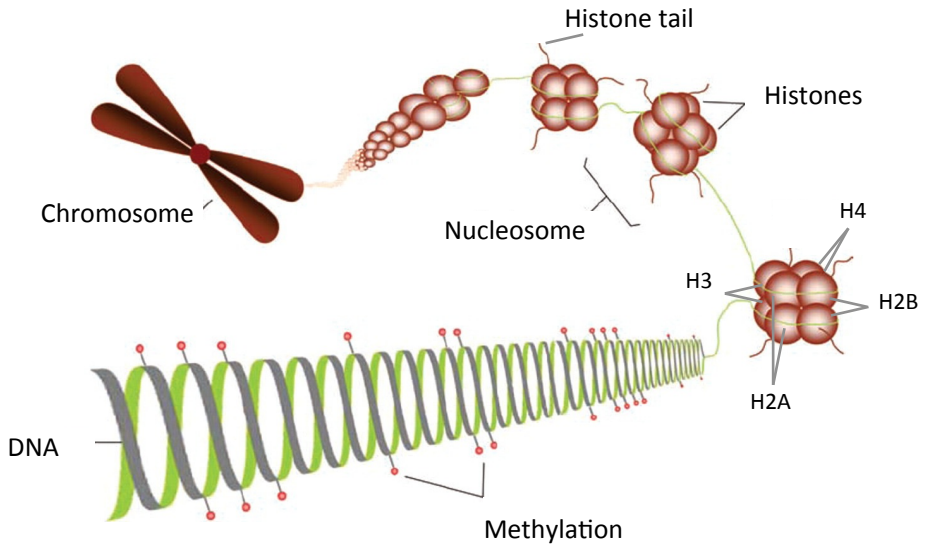


Figure 1.3 DNA and histone formation of nucleosomes [adapted from Groom et al. 2011]

Acetylation of histone 3 at lysines 9 and 14 (H3K9 and H3K14) is commonly associated with gene activation. While H3K4 and H3K36 methylation is generally associated with gene activation, whereas methylation of H3K9, H3K27 and H4K20 generally correlate with repression (Bernstein et al. 2007). Acetylation is carried out by the histone deacetylases (HDACs), which remove acetyl groups from lysine residues and thereby tighten the chromatin structure and reduce

transcription, versus histone acetylases (HATs), which perform the opposite function, adding acetyl groups to lysine residues and weakening the electrostatic charge between histones and DNA and relaxing the tightly wound chromatin (Jenuwein & Allis 2001). In essence these local changes to the chromatin structure give rise to either a relaxed (euchromatin) or condensed (heterochromatin) chromatin configuration. In a dense heterochromatic state, DNA and histone proteins are tightly packed to block the access of transcription factors and other instigators necessary for gene expression. In addition histone modifications have been found to affect neuronal function as demonstrated by the use of deacetylase inhibitors to ameliorate defects in a wide range of psychiatric and neurological conditions (Abel & Zukin 2008).

Although DNA methylation and histone modifications can act independently, they can also interact with each other. For instance, while Methyl-CpG-binding protein 2 (MeCP2) is capable of binding specifically to methylated DNA it also recruits HDACs, consequently repressing gene transcription (Nan et al. 1998). MeCP2 is also linked with histone methyltransferase activity (Fuks et al. 2003). Furthermore it has been reported that a transcriptional repression domain in DNMT1 functions by recruiting HDACs (Fuks et al. 2003), and that DNA methylation directs H3K9 dimethylation and inhibits H3K4 methylation, resulting in repression of transcription (Cedar &Malmgre Bergman 2009; Lande-Diner et al. 2007).

Thus epigenetic modifications that mark DNA sequence with regulatory information occur at multiple levels from DNA methylation, histone modification to chromatin remodelling and the displacement of nucleosomes. Though not all of these can be easily measured, epigenomic information on one form of modification can be interpreted in the context of others. For example information on open chromatin states highlighted by DNase I hypersensitivity, marks almost all active cis-regulatory elements including promoters and enhancers, but together with information on histone modification patterns that reflect the functional role of the adjoining regulatory DNA may help to delineate the precise regulatory function of a given sequence, such as the association of histone H3 lysine 4 trimethylation (H3K4me3) with promoter elements or the association of histone H3 lysine 4 monomethylation (H3K4Me1) with enhancer sequences (Thurman et al. 2012).

Integration of these different epigenomic features from publicly avaliable resources such as Encyclopaedia of DNA Elements project (The ENCODE Project, 2004) can help in defining genomic elements and while these data tracks complement each other, evidence from multiple tracks at a given sequence can be strongly suggestive of a regulatory function (The ENCODE Project, 2011). Using data from both the UCSC genome browser (human genome build hg18)

and the ENCODE project, sequences that possibly correspond to different regulatory features of the APP gene are further explored with accompanying CpG methylation data generated in Chapter 5 of this thesis.

Epigenetic regulation also extends to mechanisms involving non-coding RNAs (ncRNAs). For instance small interfering RNA (siRNA) molecules are double-stranded RNA molecules that can interfere with gene transcription and/or reshape chromatin structure. Other ncRNAs such as microRNAs (miRNAs) are tiny, highly conserved ncRNAs derived by processing of short RNA hairpins often from within promoter and regulatory sequences of the target gene itself and regulate gene expression in a post-transcriptional manner by binding to their target mRNAs, inhibiting translation or, induce the degradation of target mRNAs (Zhou et al. 2010). miRNAs are able to regulate ~30% of all human protein-coding genes by imperfect base pairing with targeting sequences and thus influence a wide variety of cellular process (Filipowicz et al. 2008) and can also regulate the expression of other types of ncRNAs (such as long ncRNAs); suggesting that these small RNAs can exert a significant impact upon transcriptomic networks (Calin et al. 2007; Zhou et al. 2010). Whereas a number of different miRNAs have been demonstrated to bind to the 3'UTR of BACE1 mRNA sequences implying their potential dysregulation in AD (Narayan & Dragunow 2010), recently an antisense RNA to BACE1; BACE1-AS transcribed from the opposite strand to BACE1 was discovered to be upregulated in patients with AD and APP transgenic mice, along with raised concentrations of Aβ (Faghihi et al. 2008). The BACE1 gene encodes, beta-site APP cleaving enzyme 1 (β-secretase) involved in the amyloidogenic APP processing pathway as described above.

1.4 Epigenetic features of AD

Despite the fact that most epigenetic studies have focussed on cancer (Jones & Baylin 2007), there are well documented epigenetic changes linked to neurodevelopmental diseases such as Rett syndrome and Fragile X syndrome (Malmgren et al. 1992; Amir et al. 1999), where the hypermethylation of the FMR1 promoter in Fragile X syndrome is directly correlated with the severity of disease (Malmgren et al. 1992). Mill et al. (2007; 2008) highlighted the role that dynamic epigenetic changes potentially play in several complex neurobehavioural diseases including major psychosis (Mill et al. 2008), attention-deficit hyperactivity disorder (Mill & Petronis 2008) and major depressive disorder (Mill & Petronis 2007); demonstrating the potential role for epigenetic mechanisms in neuropsychiatric disorders.

A number of AD features associated with a possible epigenetic role in its aetiology include: sexual dimorphism, with more women being affected than men (however this can be argued as a result of more women in older age than men; or a role of hormones, Mani & Thakur 2006); increase of clinical symptoms over time; discordant twins (Gatz et al. 1997; Nee et al. 1999; Mastroeni et al. 2009); and parental origin effect (Petronis 2001). For example LOAD families, with a strong maternal disease transmission show linkage to a region on chromosome 10 (Bassett et al. 2002; Bassett et al. 2006). Given these epidemiological, clinical and molecular peculiarities are difficult to rectify using a standard genetic approach, a multifactorial view of AD pathogenesis accommodates a combination of risk factors including epigenetic changes.

As epigenetic changes are inherited mitotically in somatic cells, they provide a possible mechanism by which the effects of external factors can produce long-term changes in phenotype. Stably acquired epigenetic changes resulting in long-lasting molecular, cellular and behavioural phenotypes, are generally though to arise either through certain environmental stimuli, particularly during sensitive periods of development (Fagiolini et al. 2009); or stochastic epigenetic changes over time (Petronis 2010). It has been suggested that stochastic epigenetic changes are more prevalent than previously thought and it is these stochastic changes in somatic cells, rather than the non-shared environment, that might account for the observed discordance between monozygotic twins (Bouchard et al. 1990; Petronis 2010). Furthermore "epigenetic drift" is proposed to encompass the age-dependent cumulative effects caused by one's environmental exposure, lifestyle, diet, drug abuse, or merely stochastic fluctuations in the epigenetic profile (Wang et al. 2008). In keeping with this, plausible DNA methylation changes in LOAD are thought to be the result of aging, diet, other environmental factors or a combination of all three.

1.5 DNA methylation & AD

1.5.1 Aging

Changes in DNA methylation have been associated with age, where aging is accompanied by a general loss of methylation (Heyn et al. 2012; Bjornsson et al. 2008). While there is a more complex pattern involving both age-related hyper- and hypomethylation (Tra et al. 2002; Christensen et al. 2009), where CpGs within CGIs show an increase in methylation, and those outside CGIs experience a decrease, similar results have been observed in human brain (Numata et al. 2012). In addition DNMT1 (which is involved in methylation maintenance) activity has been demonstrated to decrease with aging (Casillas et al. 2002) and is believed to contribute to

genomic hypomethylation during aging as a result of passive demethylation. Overall the risk of developing AD increases with age and given that age is one of the main risk factors for AD (Swerdlow 2007) accompanied changes in DNA methylation maybe related.

More specifically in relation to AD biology, an age-dependent decrease in DNA methylation status of the APP gene promoter in human post-mortem cerebral cortex (Tohgi et al. 1999a; Nagane et al. 2000) is speculated to increase transcription of this gene, resulting in overproduction of Aβ and the formation of senile plaques. Tohgi et al. (1999a) report the APP promoter to be more frequently methylated in healthy individuals between the ages of 35 and 70 years, compared those aged 74 to 90 years of age. These studies and others relating to APP promoter methylation will be discussed further in Chapter 5, where APP gene methylation is explored in post-mortem brain.

DNA methylation in the 5' region of the MAPT gene is also noted to vary with age in parietal cortex of normal control brain and was associated with an age-related decline in MAPT gene expression (Tohgi et al. 1999b). Specifically, MAPT promoter CpGs located in the Sp1 transcriptional activator binding site were hypermethylated with age, while CpGs located within the GCF transcriptional repressor binding region were hypomethylated with age (Tohgi et al. 1999b).

Siegmund et al. (2007) reported to observe accelerated age-dependent DNA methylation differences in cerebral cortex of AD cases compared to controls, with an increase in methylation status of the SORBS3 gene as well as a decrease in methylation of the S100A2 gene. SORBS3 encodes a cell adhesion molecule expressed in neurons and glia. S100A2 is a member of the S100 family of calcium binding proteins. Both these methylation changes are observed in normal aging, but in the case of AD patients, these changes were more pronounced than in normal elderly controls (Siegmund et al. 2007).

Another study exploring methylation differences in the promoters of several genes involved in AD pathogenesis and DNA methylation, observed methylation drift and unusual methylation patterns in LOAD brains relative to healthy controls in support of a role for age-dependent epigenetic drift in LOAD (Wang et al. 2008). Wang et al. (2008) found that the epigenetic distance from the norm (the median methylation of the healthy control individuals) was higher in LOAD prefrontal cortex than in healthy controls and that this “epigenetic distance” increased with aging. The promoter of TFAM, a gene responsible for mitochondrial DNA copy number regulation, was found to be most significantly affected in LOAD. While specific loci within the

promoter regions of the APOE and MTHFR genes in LOAD patients were hypermethylated compared to controls; genes that play central roles in Aβ processing (i.e. PSEN1 and APOE) and methylation homeostasis (i.e. MTHFR and DNMT1) also show a significant interindividual epigenetic variability, which may contribute to AD predisposition (Wang et al. 2008). The MTHFR gene encodes the enzyme, methylenetetrahydro-folate reductase involved in one-carbon metabolism.

1.5.2 Diet

Epigenetics could also play a role in AD through diet, a frequent observation in people with AD, is that they have low levels of folate, vitamin B12 and B6, in contrast to elevated levels of homocysteine in blood (Haan et al. 2007; Kronenberg et al. 2009). Epidemiological studies show hyperhomocysteinemia to be a possible risk factor for AD (Seshadri et al. 2002; Luchsinger & Mayeux 2004, Ravaglia et al. 2005; Clarke et al. 1998; Corder et al. 2007). For example, AD patients in the longitudinal OPTIMA (Oxford Project to Investigate Memory and Ageing) study were shown to have elevated serum homocysteine relative to cognitively normal control subjects, while individuals with highest homocysteine levels at baseline also demonstrated greatest disease progression (Clarke et al. 1998; Corder et al. 2007). Folate and vitamin B12 are required for the conversion of homocysteine to methionine and the formation of *S*-adenosylmethionine (SAM). SAM participates in biological methylation reactions, which generate *S*-adenosylhomocysteine (SAH) that subsequently forms homocysteine. Figure 1.4 shows a simplified schematic of one-carbon metabolism in relation to DNA methylation.

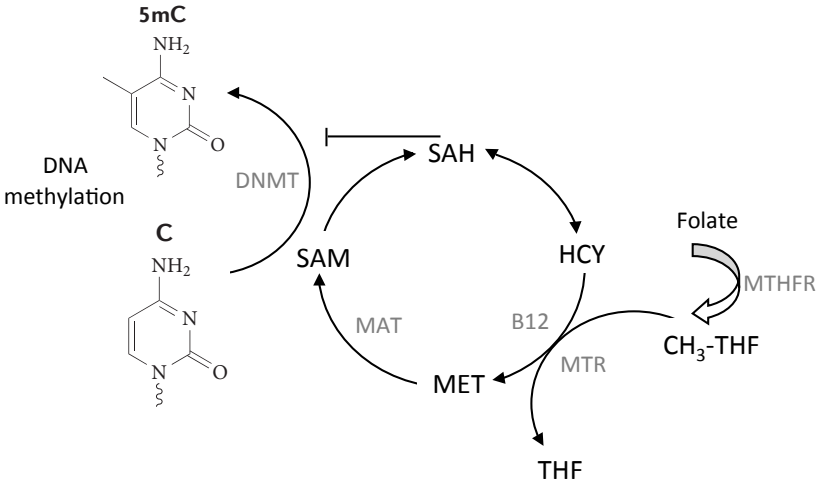


Figure 1.4 A simplified schematic of one-carbon metabolism and DNA methylation [using folate as a substrate methylenetetrahydrofolate reductase (MTHFR) forms 5-methyltetrahydrofolate (CH_3-THF). Subsequently, methionine synthase (MTR; using vitamin B12 as a cofactor) transfers a methyl group from CH_3-THF to homocysteine (HCY) forming methionine (MET) and tetrahydrofolate (THF). MET is then converted to S-adenosylmethionine (SAM) in a reaction catalysed by

methionine adenosyltransferase (MAT). Most of the SAM generated is used in DNA methylation, carried out by DNMTs, whereby SAM is converted to S-adenosylhomocysteine (SAH) by transferring the methyl group to cytosine to form 5-methylcytosine (5mC). This reaction is dependent on the clearance of by-products, and replenishment of substrates and cofactors. The expected effects of folate/B12 deficiency include (1) impairing the conversion of HCY into MET and resulting in a low level of SAM; (2) increasing the level of HCY; (3) reversing the reaction dynamics between HCY and SAH that favours SAH synthesis. High levels of SAH inhibit DNMT activity; together, these effects will eventually lead to defective genomic methylation, whereas sufficient folate/B12 prevents this from occurring; adapted from Coppedè 2010].

Animal studies in rodents have shown the disruption of homocysteine/folate metabolism to adversely affect both the developing and adult brain (Davis & Uthus 2003), such that elevated homocysteine levels can impair DNA repair mechanisms and induce oxidative stress and apoptosis in the nervous system; whereas dietary folate supplement helps clearance of homocysteine to maintain methylation homeostasis (Kruman et al. 2002; Shea et al. 2002). Also SAH is a potent inhibitor of DNMT and can lead to genomic hypomethylation (Trasler et al. 2003).

More specifically elevated SAH in AD brain inhibits DNMTs and is related to markers of disease progression and cognitive impairment (Kennedy et al. 2004). Interestingly, Durga et al. (2007) showed that folate supplementation for 3 years can significantly improve cognitive functions that decline during aging. Additionally low vitamin B12 levels in AD patients are observed to be associated with a greater basal production of IL-6 (Politis et al. 2010), not only linking diet to inflammation in AD but raising the possibility that DNA methylation has a role in the inflammatory response. Though there are likely genetic background risk factors that may also compromise the normal one-carbon metabolism or related pathways, association studies for polymorphisms in 5, 10-methylenetetrahydrofolate reductase (MTHFR), an enzyme within this pathway, have provided mainly negative findings in relation to AD (Kageyama et al. 2008; Styczyńska et al. 2008; Schjeide et al. 2009).

Furthermore studies using both cell culture and mouse experiments demonstrate a link between nutritional alterations and epigenetic modulation of AD related genes (Fuso et al. 2005; 2007; 2008; 2011). For example vitamin B6 and B12 deprivation is shown to induce hypomethylation of the PSEN1 and BACE1 promoters leading to upregulation of these genes and Aβ deposition, whereas SAM administration reverses this effect (Fuso et al. 2007; 2008; 2011). Another similar study using mouse microglial cultures showed that the APP, PSEN1 and BACE1 genes are regulated by methylation and that the administration of SAH to culture medium lead to a hypomethylation of these genes and an increase in Aβ production (Lin et al. 2009).

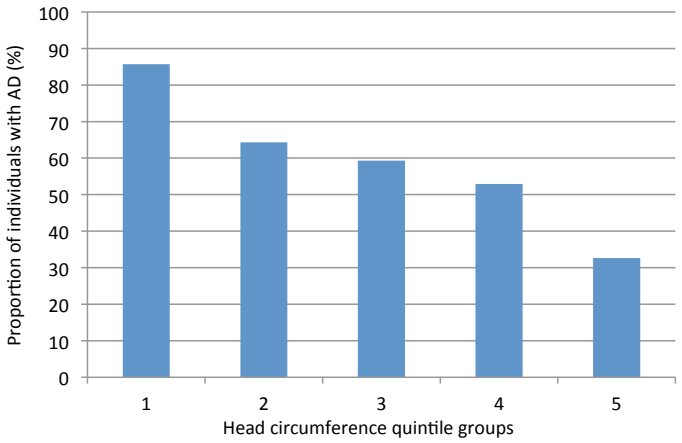
1.5.3 Environment

As mentioned before, other environmental risk factors for AD besides diet include exposure to metals or pesticides and even psychosocial factors such as stress (Wu et al. 2008; Mastroeni et al. 2009; Hayden et al. 2010; Aznar & Knudsen 2011) and given that DNA methylation (both globally and at specific loci) has been shown to be modified by environmental stimuli, such as maternal care (Weaver et al. 2004; 2005), early life stress (Murgatroyd et al. 2009) or toxins (Dolinoy et al. 2007), a growing area of research has focused on the interaction between the genome and environmental insults that may possibly be mediated by epigenetic mechanisms leading to AD later in life (Lahiri & Maloney 2010). This provides a potential mechanism through which the environment ultimately affects gene expression profiles leading to a abnormal transcriptome associated with disease.

Early life environment and its role in growth and development play an increasingly recognised role in the aetiology of AD. The Barker hypothesis, states that many adult diseases might have a developmental origin resulting from protection of key organs such as the brain in an undernourished fetus. Such changes may permanently affect the body's structure and function and lead to adult-onset diabetes, hypertension, obesity and cardiovascular disease (Barker 2006; Barker et al. 2009; Borenstein et al. 2006; Gluckman et al. 2007; Miller & O'Callaghan 2008; Dover 2009). A low-birth-weight followed by a rapid postnatal catch-up growth appears to set the stage for later glucose mishandling and obesity (Miller & O'Callaghan 2008).

Moreover obesity and diabetes are associated with a several fold increase in AD risk and may be linked to a unique form of insulin deficiency in the brain (Profenno et al. 2010).

Beeri et al. (2005) describe a relationship between body height and dementia and as height is associated with childhood nutrition it may be associated with other risk factors for dementia. As another example, smaller adult head circumference (Figure 1.5) is a recognised risk factors for AD (Schofield et al. 1997; Moceri et al. 2000; Mortimer et al. 2003; Mortimer et al. 2008; Borenstein et al. 2001; Espinosa et al. 2006; Lee et al. 2009; Wolf et al. 2004). More recently brain weight in males was correlated with DNA methylation at the IGF2 gene locus (Pidsley et al. 2009).



AD	48	36	51	18	15
Ctrl	8	20	35	16	31

Figure 1.5 Adult head circumference and AD risk
[quintile bar chart of adult head circumference and occurrence of AD; 1-lowest quintile and 5-highest quintile; N=278, normal ctrl=110, AD=168, n per group/quintile tabulated at bottom of figure; individuals with a head circumference in the lowest quintile (1) had an odds ratio of 3.6 (p = 0.003) to be affected with AD compared to individuals in the four higher quintiles; personal analysis (logistic regression, controlling for sex, age and ApoE genotype) based on data from the Alzheimer's Research Trust funded cohort at King's College London]

Environmental factors such as psychosocial stress and maternal care are shown to have marked influences on brain function through epigenetic mechanisms (Weaver et al. 2004; Murgatroyd et al. 2009). Observed methylation differences in the brains of suicide cases with childhood abuse compared to controls has also been observed (McGowan et al. 2008), further suggest a role for DNA methylation in mediating the effects of psychosocial stress factors on neuro-psychiatric phenotypes. Moreover it has been proposed that stress via epigenetic mechanisms may determine neurodegenerative disease risk (Johnstone & Baylin 2010; Kinnally et al. 2011; Babenko et al. 2012). The possible role of stressful life events and DNA methylation differences in AD associated with depression will be further discussed and explored in Chapter 4.

Animal studies have investigated the relationship between lead (Pb) exposure and epigenetics. Early life exposure to Pb in monkeys causes dysregulation of biological pathways important to LOAD pathogenesis in late life (Wu et al. 2008) and is associated with reduced DNMT activity (Wu et al. 2008; Bihaqi et al. 2011). Rat PC12 cells exposed to Pb show dose dependent decreases in global methylation and decreases in methylation at four CpG sites in the APP promoter (Li et al. 2012). These changes were associated with increases in APP mRNA and Aβ levels (Li et al. 2012).

Early life exposure of new born rodents and monkeys to low levels of Pb resulted in a late life over expression of APP, BACE1 and the transcription factor Sp1 resulting in increased levels of

Aβ peptide (Basha et al. 2005; Wu et al. 2008; Zawia et al. 2009). Rats exposed to Pb from post natal day 1 through to day 20 experienced a transient increase in APP mRNA expression in cortical brain tissue, which returned to basal levels at 1 year, and later resurged at 20 months of age in the absence of continued exposure (Basha et al. 2005). The observed late-life rise in APP mRNA was accompanied by elevated Aβ, suggesting that early life Pb exposure may have long-term effects on amyloidogenesis in late life (Basha et al. 2005). Furthermore Pb exposure is associated with cognitive decline and decreasing scores on the Mini-Mental State Exam (Wright et al. 2003; Weuve et al. 2003; Bandeen-Roche et al. 2009).

Whereas these findings provide suggestions regarding the potential link between Pb exposure and neurodegeneration, one study examining past occupational exposure to Pb (by measuring Pb levels in tibia bone in 529 former employees of a chemical manufacturing plant) suggests Pb levels and neurobehavioral test scores are influenced by APOE genotype (Stewart et al. 2002). Whereby, persistent brain effects of Pb are more toxic in subjects with at least one APOE ε4 allele (Stewart et al. 2002). The connection between past adult Pb exposure and neurodegeneration was additionally confirmed by the same group using brain magnetic resonance imaging (Stewart et al. 2006) showing an association between Pb exposure and smaller total brain volume. While these studies centred on adult occupational Pb exposure, it is not known if such workers had been previously exposed to Pb as children.

While the precise mechanism as to how metal exposure in early-life affects latent gene expression is unknown. Such environmental exposure is thought to result in oxidative damage of guanine (at CpG dinucleotides) resulting in 8-oxodG (Zawia et al. 2009). Zawia et al. (2009) note that while oxidation is a dynamic process that can occur at any time there is oxidative stress, it is typically high early and late in life, because of high metabolic rates in the former and reduced antioxidant defences in the later. The levels of antioxidant enzymes are reported to be altered in AD brain (Halliwell 2006) while other studies note increased levels of oxidative stress markers in AD brain compared to age-matched by controls (Zhu et al. 2005; Sultana & Butterfield 2010).

Though elevations in the oxidative DNA marker 8-oxodG in older animals that had been developmentally exposed to Pb (Bolin et al. 2006; Wu et al. 2008), support the idea that Pb exposure leads to latent build-up of oxidized DNA; increases in Aβ could also lead to the generation of reactive oxygen species promoting the formation of 8-oxodG. Further still the methylation pattern of cytosines could interfere with the repair of adjacent oxidized guanines (Evans & Cooke 2004). Likewise, oxidation of guanine to 8-oxodG inhibits adjacent cytosine

methylation (Weitzman et al. 1994). Also experiments using synthetic DNA oligonucleotides have shown oxidation of guanine at CpG dinucleotides greatly reduces the binding affinity to MBD proteins to these sites (Valinluck et al. 2004). Thus the synergistic effect of metal exposure and oxidative stress could possibly alter the DNA methylation landscape leading to altered gene expression, while *in vivo* animal studies support an integrated role of Pb exposure and DNA methylation in amyloidogenesis.

The effect of toxin exposure leading to global hypomethylation in human post-mortem brain has been demonstrated in a discordant AD monozygotic twin pair (Mastroeni et al. 2009). Immunohistochemistry evaluation for two markers of DNA methylation and eight methylation maintenance factors in temporal neocortex (a region exhibiting substantial AD pathology) showed a marked reduction of DNA methylation in the AD twin relative to the unaffected twin, whereas the cerebella (a brain region largely spared of AD pathology) of both twins exhibited virtually identical staining patterns and intensity for DNA methylation. The affected twin was reported to have had years of occupational exposure to pesticides that may have possibly contributed to the observed difference in methylation status at the temporal neocortex region (Mastroeni et al. 2009). These findings further support the hypothesis that the effects of environmental exposure through epigenetic mechanisms may alter AD risk and possibly provide a potential explanation for AD twin discordance despite genetic similarities (Mastroeni et al. 2009).

1.6 DNA methylation in AD post-mortem brain

Mastroeni et al. (2010) also reported to observe a similar global decrease in DNA methylation within entorhinal cortex of AD patients relative to controls. Nuclear immunoreactivity for DNMT1 and six different components of the MeCP1/MBD2 methylation complex was significantly diminished in entorhinal cortex layer II neurons of AD compared to controls (Mastroeni et al. 2010). It is suggested that these methylation differences may provide a rationale for the consistent epidemiologic and neuropathologic association of AD with homocysteine elevation and folate deficiencies (Seshadri et al. 2002; Luchsinger & Mayeux, 2004, Ravaglia et al. 2005; Clarke et al. 1998; Corder et al. 2007), since folate ultimately provides the methyl group for DNA methylation (Mastroeni et al. 2009).

Another study to investigate potential global DNA methylation differences in AD without any specific risk priori, reported AD frontal cortex brain to show to global hypermethylation relative to controls (Rao et al. 2012). Though these two latter studies report AD brain global DNA

methylation changes in opposite directions, differences are likely due to the fact that they studied different brain regions and employ different methylation detection techniques, with the latter study opting for an Enzyme-Linked Immuno Sorbent Assay (ELISA). Differences between methods for capturing global levels of DNA methylation will be further reviewed in Chapter 3.

Rao et al. (2012) also report gene specific differences in the same samples showing hyper- and hypomethylated gene promoters for the cAMP response element-binding protein (CREB) gene and nuclear transcription factor kappa B (NFKB1) genes, respectively (Rao et al. 2012). CREB has a well-documented role in neuronal plasticity and long-term memory formation in the brain (Sakamoto et al. 2011; Benito & Barco 2010). An earlier study examining AD prefrontal cortex has also shown hypomethylation relative to controls at the promoter of the CREB5 gene, which also belongs to the CRE (cAMP response element)-binding protein family (Zukin 2009). Inappropriate activation of NFKB1 has been associated with a number of inflammatory diseases and hypomethylation of NFKB1 in the AD cortex may explain reported increased neuroinflammation (Dyall 2010) due to upregulated NFKB1 activity associated with its reduced methylation state (Rao et al. 2012).

DNA methylation studies in AD include those that have studied peripheral lymphocytes (Silva et al. 2008; Wang et al. 2008) and while DNA methylation changes in blood may serve as potential biomarker of disease, the principle tissue of interest is the brain. As stated before, epigenetic profiles can differ significantly across different cell types and thus it is important to assess the organ and tissues relating to disease phenotype (Petronis 2010). The primary source of human brain for epigenetic studies is post-mortem tissue and while it is expected that DNA methylation profiles would be robust to post-mortem factors given the known stability of genomic DNA, there have been concerns relating to post-mortem interval (PMI) and its effect on DNA methylation.

To measure the effects of PMI on the feasibility of DNA methylation detection, Barrachina & Ferrer (2009) artificially increased delays up to 48 hours in the frontal cortex and hippocampus of 5 different subjects obtained after a short PMI and tested these samples alongside a standard series of case-control samples for promoter DNA methylation differences in a number of AD and tauopathy-related genes (ADORA2A, APP, MAPT, PSEN1, RAGE and UCHL1). Using a sensitive DNA methylation analysis method based on matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) no differences in the percentage of CpG methylation were found between control and disease samples analysed, nor did they note any difference in the test loci for artificially increased PMI of up to 48 hours (Barrachina & Ferrer 2009). While these results

found no difference in DNA methylation at the promoters of the genes tested for AD cases and controls the appraisal of PMI on DNA methylation detection indicate that there is likely to be no effect of PMIs of up to 48 hours on DNA methylation stability.

To date most of the promoter methylation gene studies in AD post-mortem brain have investigated a handful of candidate genes at once, more recent studies take a unbiased approach to genome-wide discovery of locations of DNA methylation differences, which allows for identification of novel disease-associated genes. For example Bakulski et al. (2012) used the Illumina Infinium HumanMethylation27 BeadArray assay to determined DNA methylation at 27,578 CpG sites spanning 14,475 genes in human frontal cortex brain tissue between LOAD cases and controls. While analysis revealed 948 CpG sites representing 918 unique genes as potentially associated with LOAD disease status their top hit was located in the promoter of the transmembrane protein 59 (TMEM59) gene with an observed 7.3% hypomethylation in cases relative to controls. Methylation at this site was also associated with corresponding mRNA and protein levels of the TMEM59 gene product (Bakulski et al. 2012). The authors note that this change was independent from DNA methylation changes with age and that PMI was not a significant confounder at the TMEM59 site (Bakulski et al. 2012). Also the TMEM59 gene was previously implicated in APP protein post-translational processing, supporting a role for epigenetic change in LOAD pathology (Ullrich et al. 2010).

Within the study by Bakulski et al. (2012), looking at genes known to be associated with AD, there was one CpG site representing PSEN1 on the array that associated with a 1% decrease in methylation for AD samples compared to controls (Bakulski et al. 2012).

In summary various lines of evidence point to DNA methylation as having a possible role in the aging brain and AD (Table 1.1). From observations of global methylation differences in AD case-controls; a case study of discordant MZ twins (Mastroeni et al. 2010; 2009); methylation differences highlighted at the promoter of the APP gene (West et al. 1995; Tohgi et al. 1999a; Basha et al. 2005; Wu et al. 2008; Lin et al. 2009; Li et al. 2012); to those in genes known to influence Aβ production via the amyloidogenic pathway (Scarpa et al. 2003; Fuso et al. 2007; 2009; 2011; Wu et al. 2008; Lin et al. 2009; Bakulski et al. 2012). Together these latter changes increase the levels of Aβ which in turn give rise to the formation of senile plaques. Other tentative gene promoter methylation differences in AD cases-controls include the APOE, MTHFR, TAMF, SORBS3, S100A2 CREB, NFKB1, CREB5, and TMEM59 genes (Wang et al. 2008; Siegmund et al. 2007; Zukin 2009; Rao et al. 2012; Bakulski et al. 2012).

Disease relevance	Gene(s)	Direction of methylation change	Associated changes	Tissue	Reference
Age related	APP	↓	-	Human parietal cortex	Tohgi et al. 1999a
	MAPT	↓↑	-		Tohgi et al. 1999b
Age / AD related	SORBS3	↑	-	Human cerebral cortex	Siegmund et al. 2007
	S100A2	↓	-		
	APOE MTHFR TFAM	↑	-	Human prefrontal cortex	Wang et al. 2008
AD related	APP	↓	-	Human temporal lobe	West et al. 1995
	Global	↓	DNMT1 and six different components of the MeCP1/MBD2 methylation complex diminished	Human entrorhinal cortex	Mastroeni et al. 2010
	Global	↑	-	Human frontal cortex	Rao et al. 2012
	CREB	↑	-		
	NFKB1	↓	-		
	CREB5	↓	-	Human prefrontal cortex	Zukin 2009
	TMEM59 PSEN1	↓	Increased TMEM59 mRNA and protein levels	Human frontal cortex	Bakulski et al. 2012
AD related / toxin exposure	Global	↓	-	Human temporal neocortex	Mastroeni et al. 2009
Pb exposure	APP BACE1 Sp1	-	Late-life over expression of APP, BACE1 and Sp1, accompanied by elevated Aβ	Long-Evans hooded rat strain brain early life exposure to Pb	Basha et al. 2005
	APP	↓	Global hypomethylation, increases in APP mRNA and Aβ levels	Rat PC12 cells exposed to Pb	Li et al. 2012
	APP BACE1 Sp1	-	Late-life over expression of APP, BACE1 and Sp1, accompanied by elevated Aβ and reduced DNMT activity. oxidative DNA marker 8-oxodG	Monkey (*Macaca fascicularis*) brain early life exposure to Pb	Wu et al. 2008
Diet related	APP BACE1 PSEN1	↓	Increase in APP mRNA and Aβ production	Mouse microglial cultures after SAH administration	Lin et al. 2009
	PSEN1	↑	Downregulated PSEN1 gene expression and Aβ production	Human neuroblastoma SK-N-SH cells after SAM administration	Scarpa et al. 2003
	PSEN1 BACE1	↓	PSEN1 and BACE1 up-regulation & Aβ deposition - reversed by SAM adminstration	Human neuroblastoma SK-N-BE cells; TgCRND8 & 129Sv mouse strain brain; with induced vitamin B6 and B12 deficiency	Fuso et al. 2007; 2008; 2011
	PSEN1 BACE1	No change		Human A172 glioblastoma cells with induced vitamin B6 and B12 deficiency	Fuso et al. 2007

Table 1.1 Summary of cited methylation changes relating to AD

Together any number of these changes on their own or in combination at low levels could lead to impaired DNA methylation in aging individuals and thus increased risk of developing AD. Whether it be a global age associated shift or a specific DNA methylation change, epigenetics has the potential to explain some of the risk component in LOAD. Potentially these changes could act as peripheral epigenetic biomarkers for AD. Further more the dynamic nature of the epigenome means that, unlike pathogenic DNA sequence mutations, epigenetic disruption is potentially reversible, and thus a realistic target for pharmacological intervention.

1.7 Aims of this thesis

Given that relatively few studies to date have explored the potential of DNA methylation changes in AD post-mortem brain. The overall aim of this thesis was to explore different aspects of AD risk (aging, depression and amyloid-β deposition) by examining case-control differences in DNA methylation using post-mortem tissue covering brain areas affected by AD pathology and those that are spared of AD pathology or associated with behavioural phenotypes. Both global DNA methylation and DNA methylation of two candidate genes relating to AD risk factors were examined for differentially methylated regions. The three studies presented in this thesis represent different aspects of AD risk, which include aging (explored through global DNA methylation), depression (and SERT promoter methylation) and amyloid-β deposition in brain (by exploring APP gene region methylation) the broad aims of these studies are as follows:

I. Chapter 3 explored global genome-wide DNA methylation levels using a number of different measures and techniques to ascertain if there were measureable methylation differences in AD pathology affected brain area and non-AD pathology affected areas across cases and controls. Global methylation levels were assed for age-depended changes and possible patterns of DNA demethylation relating to the conversion of 5mC to 5hmC. To achieve this a novel method for detecting global methods of 5mC and 5hmC content of DNA methylation was developed and assessed as part of this study, in doing so this gives a more accurate picture of DNA methylation differences that may underlie DNA methylation changes specific to the brain given that 5hmC is at higher levels in the brain relative to other tissues and tentatively linked to the neurodegenerative process (Song et al. 2010).

II. Chapter 4 examined the association between the serotonin transporter (SERT) gene promoter DNA methylation and mRNA abundance across seven different brain areas in relation to AD status and additionally assessed the associations of 5HTTLPR genotype, the presence of stressful life events and depression in AD samples. In combining these data to understand whether SERT mRNA expression changes in the different brain areas associate with depression in AD, it was possible to test whether or not these changes are possibly modulated by 5HTTLPR polymorphic variation, environmental stresses, and DNA methylation at the promoter.

III. Chapter 5 examined the association between APP gene region DNA methylation and AD pathology markers along with APP protein levels in an AD pathology affected brain area. The aim was to prepare a 96 sample (case-control) barcoded bisulphite PCR sequencing library using a microfluidics system and to assess its novel application to highly parallel bisulphite sequencing. While employing such a methodology provided base-pair resolution of high depth counts of individual CpG sites across the entire CGI and additionally selected sequences of the APP gene region, including those thought to be of regulatory potential as indicated by data from the ENCODE project. SNP calls from the same bisulphite sequencing data were also explored for association with nearby CpG methylation levels that may represent methylation quantitative-trait loci.

MATERIALS & METHODS

2.1 DNA extraction from frozen brain tissue

The preparation of pure samples of DNA from any sort of tissue plays a central role in molecular biology. The process involves three main steps. The first steps involve the disruption and breaking open the cells from which the DNA will be obtained, by both physical and enzymatic means. Secondly, the removal of cell debris and protein, by a denaturing agent such as phenol. Lastly the DNA is precipitated out using salt and ethanol, whereby ethanol is added to the aqueous solution, which causes the nucleic acid to precipitate out of solution. Samples were processed using a standard phenol-chloroform extraction as outlined below.

1. Make up the following reagents the day before use:
   i. Autoclaved lysis buffer: 75mM NaCl; 10mM Tris-HCl pH 8; 25mM EDTA pH 8
   ii. 10% (w/v) SDS
   iii. Autoclaved 1X TE buffer (pH 8.0): 10mM Tris-HCl (pH 8.0); 1mM EDTA
2. On the day of use prepare the lysis buffer with the addition of the SDS , by adding 1/20th the volume of 10% (w/v) SDS, resulting in a final concentration of 0.5% SDS.
3. In a 1.5 mL tube to 100 mg of frozen tissue add 225 µL lysis buffer and 25 µL proteinase-K (at 20 mg/mL). Mash up the tissue as much as possible with a sterile disposable pestle.
4. Add a further 350 µL lysis buffer, break up tissue further with vigorous pipetting using filter tips and a P1000.
5. Place the tubes in a Thermomixer Comfort heating block (Eppendorf) at 55^{o}C overnight with mixing at 750 rpm.
6. The following morning, a visual inspection is made to assess if a successful digestion of the tissue has been achieved. Fatty tissue may remain although this will not affect further sample processing.
7. Add 3uL RNase A/T1 (A - final concentration of 10 µg/mL, and T1 - final concentration of 25 Units/mL) and incubate at 37 ^{o}C for a minimum of 1 hour.
8. Incubate at 65^{o}C for 20 minutes to inactivate the proteinase-K.
9. Add 650 µL of room temperature Tris equilibrated phenol/chloroform, manually shake/invert the tubes 20 times and separate the phases by centrifugation at 13,000 rpm for 15 minutes.

10. Carefully transfer ~550 µL of the upper aqueous phase to a new labelled 1.5 mL tube.
11. Add 550 µL Chloroform, manually shake/invert the tubes 20 times and separate the phases with by centrifugation at 13,000 rpm for 15 minutes.
12. Carefully transfer ~420 µL of the upper aqueous phase to a new labelled 1.5 mL tube. Repeat this Chloroform extraction step.
13. Add (2.5X volumes) ~900 µL 100% ethanol, manually invert tubes 10-15 times. If pellet does not precipitate add additional 100 µL volumes of ethanol until it does otherwise see below*. The DNA is spooled out of solution on a sealed glass rod (or sterile plastic loop). Carefully wash with 70% ethanol (dip into 500 µL 70% ethanol in a separate tube for each sample, make sure the pellet doesn't dissociate from the rod) and the DNA is allowed to air dry at room temperature for 2-3 minutes.
14. The DNA is transferred from the tip of the rod into 200-300 µL of 1X TE in a labelled 1.5 mL tube. Dissolve overnight at 37 ^{o}C.
15. If the pellet is too small to spool out, spin the tube 10 minutes at 13,000 rpm, carefully remove supernatant. Wash with 800 µL 70% ethanol, invert 3x and spin down 5 minutes at 13,000 rpm. Carefully remove supernatant and leave pellet to air dry for 15-20 minutes and resuspend in ~70 µL 1X TE
16. Visually inspect an aliquot of the extracted DNA on a 0.8% agarose by gel electrophoresis, to ensure the extracted DNA sample is not degraded and is of high molecular weight.
17. Check the quality and concentration of the extracted DNA after gel electrophoresis by Nanodrop spectrophotometry.

*If there is very little or no DNA visible on adding additional volumes of 100% ethanol add either Ammonium Acetate to a final concentration of 2.5M, OR Sodium Acetate to a final concentration of 0.3M to aid the precipitation. Then follow the steps above for DNA that is too small to spool out.

2.2 Agarose gel electrophoresis

Agarose gel electrophoresis is routinely used to separate and visualize double stranded DNA fragments. Depending on the application the agarose matrix can vary from 0.7% to 3%. While a lower percentage gel will permit the migration of large fragments (2-25Kb) a higher percentage gel allows a more accurate separation of smaller sized fragments (0.1-2Kb) as the agarose matrix is denser. The separated products are visualized via UV transluminescence of the ethidium bromide stained DNA.

1. Make up the following reagents the day before use:
   i. 10X TBE buffer pH 8.2: Tris base 0.89M; EDTA 0.02M; Boric acid 0.89M
   ii. 6X sample loading buffer: dissolve 4g sucrose and 2.5mg bromophenol blue in 6 mL of 1X TE and bring the final volume to 10ml with 1X TE buffer.
2. The desired % w/v agarose is added to 100 mL of 1X TBE buffer.
3. Mix and then dissolve by heat in a microwave oven for ~1 minute. Heat for further bursts of 30 seconds until the agarose is completely dissolved, taking care not to let the mixture boil over as it becomes super-heated.
4. Allow the mixture to cool a little and add ~2.5 µL of ethidum bromide to a final concentration of 0.5 µg/mL.
5. Mix in well taking care not to introduce air bubbles and then cool down further to 50-60°C before pouring into a prepared gel try and appropriately sized combs.
6. Allow the get to cool and completely set before use.
7. Place the gel in a tank containing 1X TBE buffer and then top up with further 1X TBE buffer until the gel is submerged by a 2-5mm of buffer.
8. Prepare 1-5 µL of the DNA samples with loading buffer to 1X and load alongside an appropriately sized DNA ladder.
9. Run the gel at ~80-150v for ~0.5-1 hour depending on the separation/resolution required.
10. Once run, visualize the gel under UV light.

2.3 Nanodrop spectrophotometer

Routine measurement of DNA concentration was made using a Nanodrop spectrophotometer. The Nanodrop allows measurement on a very small amount of DNA, while comparison of the optical density values of the solution at various wavelengths can be used to determine the purity of the sample solution. Typically 1 µL is sufficient for this purpose, however the concentration is ideally determined by averaging the concentration of 3 separate readings on 1.5 µL of sample.

1. Launch the Nanodrop software and select ‘Nucleic Acid’ measurement.
2. Wipe both the lower and upper pedestal with a lint-free tissue and distilled water.
3. Add 1.5 µL of distilled water and select ‘Initialize’.

4. Wipe pedestals clean and add 1.5 µL of the appropriated buffer used to dilute the DNA samples that are to be tested, select 'Blank'.
5. Wipe pedestals, for each sample enter a sample ID; add 1.5 µL DNA sample and select 'Measure'.
6. Select 'Show Report', chose 'Save Report' and export as a tab-delimited text file.
7. When finished wipe both the lower and upper pedestal with a lint-free tissue and distilled water.
8. Check optical density ratios. For pure DNA, the observed 260/280 nm ratio will be near 1.8. Elevated ratios usually indicate the presence of RNA, which can be tested by running the sample, on a 1% agarose gel. 260/280 ratios below 1.8 often signal the presence of a contaminating protein or phenol. Alternatively, protein, salt or phenol contamination is indicated by 260/230 ratios greater than 2.2.

2.4 Agilent 2100 Bioanalyzer

Agilent DNA 1000 kit chips and reagents designed for sizing in the range of 25–1000bp and a quantitative range 0.1–50 ng/µL. Each chip contains an interconnected set of microchannels that is used for separation of nucleic acid fragments based on their size as they are driven through it electrophoretically.

For accurate determination of DNA concentration, DNA samples were diluted to between 0.1–50 ng/µl, if they were known to be of a particularly high concentration. Samples were tested using the following protocol.

1. Allow the gel-dye mix to equilibrate to room temperature for 30 min before use.
2. Put a new DNA chip on the chip priming station.
3. Pipette 9 µL of gel-dye mix in the well marked G within a black circle.
4. Make sure that the plunger is positioned at 1 mL and then close the chip priming station.
5. Press plunger until the clip holds it.
6. Wait for exactly 60 seconds then release clip.
7. Wait for 5 seconds and slowly pull back plunger to 1 mL position.
8. Open the chip priming station and pipette 9 µL of gel-dye mix in the wells marked with a bold G.

9. Pipette 5 µL of marker (green tube) in all 12 sample wells and ladder well. Do not leave any wells empty.
10. Pipette 1 µL of DNA ladder (yellow tube) in the well marked with a ladder.
11. In each of the 12 sample wells pipette 1 µL of sample (used wells) or 1 µL of de-ionized water (unused wells).
12. Put the chip horizontally in the adapter of the shaker and vortex for 1 min at 2400 rpm.
13. Run the chip in the Agilent 2100 Bioanalyzer within 5 min, using the DNA 1000 chip type specific parameters.
14. The chip takes roughly 35 minutes to run.
15. The run is automatically saved to a folder with the chip barcode name and can be analysed further using the Agilent 2100 Bioanalyzer software. Depending on the desired analysis various images and measures can be made using the software.

2.5 Quant-iT PicoGreen DNA quantification

PicoGreen flurometry is used instead of the Nanodrop spectrophotometer to accurately quantify DNA as the dye is DNA specific and readings are less likely to be affected with other contaminants. PicoGreen dsDNA quantitation reagent is an ultra sensitive florescent nucleic acid stain for quantitating double stranded DNA in solution. It does not detect common contaminants such as proteins, RNA, ethanol, salts, and ssDNA

1. Make up a dilution of the DNA samples according to the current known concentrations to 0.5 ng/µL in a volume of 50 µL (e.g. 1 µL of DNA at 25 ng/µL in 49 µL water). For greater accuracy dilute/measure each sample in triplicate. Plate the dilutions into a cliniplate
2. For the control sample (λ DNA) calibration curve make a 1:100 dilution to 1 ng/µL in water and then six further serial dilutions (0.5, 0.25, 0.125, 0.0625, 0.03125 and 0.015625) by adding 50 µL of water and then taking 50 µL of that for the next dilution. For the final calibration sample use 50 µL of water, resulting in 8 calibration samples. Again plate these into a single column of a cliniplate.
3. Make up a fresh PicoGreen solution at a 1:200 dilution. Prepare 50 µL per sample (0.25 x (N + 8 for standard curve + 8 for overhang)). Protect the working dilution from light by wrapping it in foil, as it is susceptible to photodegradation.
4. Switch on the Ascent fluorometer machine and open the Acent software from the PC desktop. Use the default settings.

5. To run the calibration samples first. In the 'Plate Layout' tab set the 'Area Definition' - highlight the wells to be tested and for the 'Layout' - specify sample information for each well.
6. From the dropdown menu select 'Well' then 'Fill' then 'Calibr' tick 'Generate Conc. Series'. Specify the 'Number of Calib' as 8 and the 'Divide By' option as 2, then click 'Apply' and 'Close'.
7. At this stage using a multichannel pipette add 50 µL PicoGreen to each of the calibration dilutions then insert the plate, and click 'Start'. To save the file once it has completed select the 'Results' tab then 'Sheet' and 'Save As'.
8. Check the curve fit. Open the saved file select 'Processes' from the dropdown menu then 'Curve Fit', then 'Source' and choose the saved file as both the calibrator source and the sample source. Check that the curve is linear and the correlation coefficient is ~0.9.
9. To run the samples select the 'Procedure' tab, 'General' button and set the 'Area Definition' - unselect previously defined control sample wells and reselect wells for test samples. Set the 'Layout' - specify sample information for each well, by selecting 'Fill' then 'Sample' from the dropdown menu, use the default settings specifying the number of samples. click 'Apply' and 'Close'.
10. Using a multichannel pipette add 50 µL PicoGreen to each of the test sample wells. insert the plate, and click 'Start'. To save the file once it has completed select the 'Results' tab then 'Sheet' and 'Save As'.
11. Select the 'Results' tab, open the saved test sample file, use the 'Curve' function to run the samples against the saved calibration file from earlier. To save a copy of these results, copy and paste them into an Excel spreadsheet.

2.6 Sodium bisulphite treatment of DNA samples

The ability to detect and quantify DNA methylation via PCR based techniques requires the starting template to be converted such that the methylated cytosines can be distinguished from unmethylated cytosines. The most commonly used method to do this is bisulphite conversion. Whereby all unmethylated cytosine residues are converted into uracil during the bisulphite treatment. These cytosines (C) are then treated as thymines (T) for all downstream PCR based applications and primer design purposes. The analysis of such converted sequences is therefore reduced to differentiating between the chemically induced polymorphisms (cytosines and thymidine) resulting from bisulphite conversion.

The recovered DNA is typically A, U, and T-rich and the original base-pairing no longer exists. Instead, it is single stranded with limited non-specific base-pairing at room temperature. The absorption coefficient at 260 nm resembles that of RNA. Use a value of 40 µg/ml for Ab260 = 1.0 when determining the concentration of the recovered bisulphite-treated DNA. Note such measurements made using the Nanodrop only provide an indication of concentration and are not to be taken as absolute.

In general two different methods were used for bisulphite treatment depending on the number of samples that were to be processed. Smaller numbers were converted using the Microcon YM-50 column method while EZ-96 D5003 shallow-well plates were used for a larger number of samples.

2.6.1 Millipore Microcon YM-50 columns method

1. Prepare the following reagents on the day of use:
   i. Sterile double-distilled water, preferably freshly degassed under a vacuum (free oxygen in water can reduce the efficiency of sodium bisulphite conversion), or alternatively use a DNase, RNase free PCR grade water.
   ii. Fresh 3M NaOH solution - dissolve 3 g NaOH pellets in 25 mL water.
   iii. Fresh 0.1M NaOH solution made from a dilution of above.
   iv. Fresh hydroquinone solution. Dissolve 0.22 g hydroquinone in 10 mL water. Keep this solution shielded from light.
   v. Saturated sodium bisulphite solution. Bring 10.8 g sodium bisulphite to 16 mL final volume in preheated water (55°C). Invert to mix until solution is fully saturated. Add 2.6 mL 3M NaOH solution and 1.0 mL hydroquinone solution. Mix well. It is possible that the sodium bisulphite will not entirely dissolve. If any substrate remains, centrifuge solution before use and use the supernatant.
2. Take 500ng of the DNA sample(s) and adjust volume to 10 µL. Transfer DNA sample(s) to PCR tubes (or 96-well plate for high-throughput processing). Add 1.1 µL of fresh 3M NaOH solution. Spin down and seal tubes (or plate).
3. Place in a thermocycler for 20 minutes at 42˚C.
4. Spin down tubes/plate to catch condensation and carefully open seal. Add 120 µL of fresh sodium bisulphite solution, seal plate/tube with a fresh lid, invert a few times to mix, and then spin-down.
5. Place in a thermocycler for 4-5 hours at 55˚C. It can be beneficial to ramp the reaction up to 95°C for 1 minute every hour to ensure that the DNA remains single-stranded.

6. Remove plate/tubes from thermocycler, spin down, and carefully remove lid. Add 100 µL of sterile water.
7. Transfer each sample to a Microcon YM-50 column (Millipore) and draw solution through filtration matrix by centrifugation at maximum for ~10 minutes, until wells are visibly empty.
8. Desalt DNA by adding 175 µL of sterile water to each well and draw the solution through the matrix via centrifugation as before for ~4 minutes. Repeat this step another two times.
9. Desulfonate by adding 175 µL fresh of 0.1M NaOH and draw the solution through the matrix via centrifugation for ~4 minutes.
10. Perform a final wash step by drawing 175 µL of sterile water through the matrix with centrifugation for ~4 minutes.
11. Recover DNA by adding 30 µL of 1 x TE Buffer and incubate for 2 min at room temperature.
12. Carefully vortex the uncapped Microcon column for additional 30 seconds, separate sample reservoir from filtrate cup and place sample reservoir upside down into a new 1.5 mL tube. Spin for 3 minutes at 1780 x *g* (4500 rpm) in invert spin mode to elute DNA. This should roughly give a bisulphite converted DNA template of ~15 ng/µL.
13. Remove an aliquot for immediate use if required. Otherwise, store bisulphite treated DNA at -20˚C (or -80˚C for long-term storage). It is preferable to make converted sample aliquots of 2-3 µL in 0.2 mL strip-tubes or 96-well plates for later use in PCR.

2.6.2 EZ-96 D5003 shallow-well format Zymo Research plates

The EZ DNA Methylation™ Kit uses a simplified procedure and streamlines the bisulphite method for DNA methylation analysis in a 96-well plate format while the Kit's innovative column desulphonation reaction eliminates several precipitation steps, reduces template degradation and minimizes DNA loss during treatment and clean up.

1. The CT Conversion Reagent supplied within this kit is a solid mixture and should be prepared on the day of use. Add 7.5 mL water (preferably degassed) and 2.1 ml of M-Dilution Buffer to a bottle of CT Conversion Reagent. Mix at room temperature with frequent vortexing or shaking for 10 minutes. It is normal to see trace amounts of undissolved reagent in the CT Conversion Reagent. The CT Conversion Reagent is light sensitive; so minimize its exposure to light. For best results, the CT Conversion Reagent should be used immediately following preparation.

2. Add 5 µL of M-Dilution Buffer to each DNA sample in a Conversion Plate and adjust the total volume of 50 µL with water. Mix each sample by pipetting up and down.
3. Incubate the Conversion Plate containing the samples at 42°C for 20 minutes.
4. After the above incubation, add 100 µL of the prepared CT conversion Reagent to each sample and mix.
5. Incubate the Conversion Plate in the dark at 50°C for 12-16 hours (using a thermal cycler). Note: it can be beneficial to ramp the reaction up to 95°C for 1 minute every hour to ensure that the DNA remains single-stranded. Since the CT conversion reagent is light sensitive try to minimize the reaction's exposure to light whenever possible.
6. Incubate the sample at 2°C (e.g. on ice or using the thermal cycler) for 10 minutes
7. Add 400 µL of M-Binding Buffer to each well of a Silicon-A Binding Plate on a Collection Plate.
8. Load the samples from the incubation step into the wells of the Silicon-A Binding Plate containing the M-Binding Buffer. Mix by pipetting up and down.
9. Centrifuge at ≥ 3,000 x *g* (5,000 x *g* max.) for 5 minutes. Discard the flow through.
10. Add 500 µL of M-Wash Buffer to each well and centrifuge at ≥ 3,000 x *g* for 5 minutes.
11. Add 200 µL of M-Desulphonation Buffer to each well and let stand at room temperature (20-30 °C) for 15-20 minutes. After the incubation, centrifuge at ≥ 3,000 x *g* for 5 minutes.
12. Add 500 µL of M-Wash Buffer to each well and centrifuge at ≥ 3,000 x *g* for 5 minutes. Add another 500 µL of M-Wash Buffer and centrifuge for 10 minutes. It is important that the ethanol in the M-Wash Buffer is efficiently removed from the matrix by either centrifuge speed or drying. Even if no M-Wash Buffer is visible in the well and the matrix is still holding residual buffer, it will interfere with elution and downstream applications. If it is not possible to use a higher centrifuge speed, the final wash step should be lengthened to at least 20 minutes or the plate should be dried at 37°C for 30 minutes prior to elution.
13. Place the Silicon-A Binding Plate onto an Elution Plate. Add 30 uL of M-Elution Buffer directly to the binding matrix in each well. Centrifuge for 3 minutes at ≥ 3,000 x *g* to elute the DNA.
14. The DNA is ready for immediate analysis or can be stored at -20°C for later use. For long term storage, store at -70°C. For each PCR, 1-4 uL of eluted DNA is recommended.

- Additional notes on column recovery:

To aid sample recovery adhere to the following: (i) make sure that the elution buffer is added directly to the column matrix, not the side of the well; (ii) allowing the buffer to sit on the column for several minutes prior to centrifuging. Also elution buffer that has been warmed to 50-60^{o}C can also be used to aid sample recovery.

Zymo do not recommend using less than a 30 µL elution volume with the shallow well Silicon-A Plates. The problem being that the matrix will not be fully saturated by the elution buffer and recovery of the converted DNA is likely to be incomplete and inconsistent. Elution of smaller volumes such as 15 µL should be carried out in the deep-well Zymo-Spin I-96 Plates (D5004).

2.7 Bisulphite converted DNA PCR design and optimisation

The choice of CpG island sequences for primer design was performed using the various web based genome browsers (UCSC, Ensembl and NCBI) and their annotated data tracts for regulatory features. Then sequences were identified and extracted for the promoter/CGI region along with known SNPs.

Where possible the literature characterizing a particular gene and promoter region was consulted to either confirm or further identify CGI sequences, any specific binding sites or marked regulatory areas in the 5' region/promoter/CGI that could be of potential interest and associated with regulation of the gene. Where such sequences were identified they were included within designed PCR amplicons for the give gene.

Bisulphite PCR primers were designed using the MethPrimer (Li & Dahiya 2002) web tool, using the default parameters. Known SNPs were highlighted (by arrow brackets) so as to exclude them from the primer sequences. After designing potential primers the BiSearch (Tansady et al. 2005) web tool was used to check for non-specific primer binding to non-target regions of the bisulphite converted genome. Each primer is searched on the two non-complementary genomes created by bisulphite transformation of the original sense and anti-sense strands. In addition the primers are also searched on the two additional virtual strands that only present after the first few PCR cycles Figure 2.1.

Primers were then ordered with any additional assay specific requirements such as tag sequences or purification. Using the guidelines in the table below, new primer pairs were

optimised for PCR using a previously amplified bisulphite treated DNA as template along with a 100% and 0% methylated converted samples. PCR products were then visualised on a 2% agarose gel.

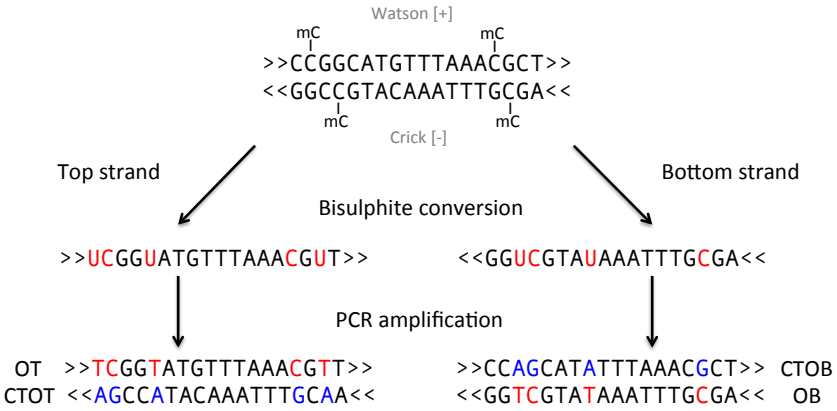


Figure 2.1 Illustration of bisulphite conversion & PCR
[bisulphite conversion of genomic DNA gives rise to two uncomplimentary strands denoted as the top strand (+) and bottom strand (-), methylated cytosines are resistant to bisulphite conversion, whereas unmethylated cytosines are converted to uracil (U); subsequent PCR amplification can produce up to four potentially different DNA fragments for any given locus; mC: 5-methylcytosine; OT: original top strand; CTOT: strand complementary to the original top strand; OB: original bottom strand; CTOB: strand complementary to the original bottom strand; adapted from Krueger et al. 2012]

PCR parameter	Considerations for PCR optimisation
Number of cycles	35-40 cycles are required for successful PCR amplification of bisulphite converted DNA.
Annealing temperature	55-60°C typically work well. As the treated DNA is AT-rich and has low GC composition, it may be necessary to reduce the annealing temperature accordingly.
DNA Polymerase choice	A hot-start polymerase (e.g. Qiagen HotStar) is used to reduce non-specific amplification, which is relatively common with bisulphite treated DNA due to its AT-rich nature.
DNA Polymerase concentration	0.1U per reaction works for most PCRs but some difficult PCRs may require up to 0.4U.
Primer concentration	20-50nM low final concentrations help reduce primer-dimer and non-specific amplification.
MgCl concentration	3-5mM final concentrations at the higher end can aid difficult bisulphite PCR. Note that the 1X PCR buffer for Taq polymerase Qiagen HotStar (Taq/buffer system used throughout the thesis) already has a final MgCl of 1.5mM.
Converted template	Avoiding or reducing the number of freeze-thaw cycles the bisulphite converted template goes through is linked to greater success in PCR.

2.8 Capillary electrophoresis Sanger sequencing

1. Check 1-2 µL of PCR products on a 2% agarose gel to roughly quantify the amount of product generated. Dilute the products to ~2.5 ng/µL for further clean up and sequencing. As a rough guide, dilute a good strong PCR product to 1 in 10. Products are

then cleaned up to remove unwanted PCR components and by products using microCLEAN solution.

2. Add an equal volume of microCLEAN to PCR product. Mix by pipetting and leave at room temperature for 5 minutes. If a small number of samples are to be cleaned up tubes can be used, whereas a larger number of samples are to be processed use a 96-well plate as outlined below.
3. For a clean up in tubes, spin at 13,000 rpm in microfuge for 7 minutes. Remove supernatant and spin briefly again and remove all remaining liquid. Resuspend pellet in the appropriate volume of 1X TE and leave for 5 minutes to rehydrate the product.
4. For a clean up in a 96-well plate plates, spin at 2000-4000 g for 40 minutes. Place the plate upside down onto tissue paper in the centrifuge holder, pulse centrifuge to <40 g for 30 seconds. Resuspend pellet in the appropriate volume of 1X TE and leave for 5 minutes to allow product to rehydrate.
5. The clean PCR product can then used as template for dye-terminator Sanger sequencing. Sequencing reactions are set up and cycled as outlined below.

Component	Final concentration
Big Dye Diluent [5X]	0.75X
Big Dye v3.1 [2.5x]	0.25X
Primer Fwd or Rev	0.16µM
Template	5 ng
PCR grade water	To a final volume of 10µL

Big Dye sequencing reaction mix

Step	Temperature	Time
Initial	96°C	1 min
Denaturing	96°C	10 sec
Annealing	50°C	5 sec
Extension	60°C	4 min
Final	10°C	3 min
25 Cycles in a dyad PCR machine		

Cycling conditions

6. Sequencing reactions are cleaned up in a 96-well plate using ethanol precipitation as follows.
7. Add 2.5 µL of 125 mM EDTA to each well, tap the plate to ensure the EDTA reaches the bottom of the wells.
8. Add 30 µL 100% ethanol, seal the plate and mix by inverting the plate 4 times. Incubate at room temperature for 15 minutes.

9. Then centrifuged at 1500 x *g* at 4°C for 45 minutes. Remove the plate seal and invert the plate onto a paper towel, centrifuge at 185 x *g* for 30 seconds.
10. Remove from the centrifuge and add 30 µL 70% ethanol and centrifuged at 1500 x *g* at 4°C for 15 minutes. Remove the ethanol by a short inverted spin for 1 minute as before.
11. Briefly air dry the samples for 5 minutes and immediately setup for capillary electrophoresis on an ABI 3100 sequencer.
12. Resuspended the samples in 10 µL of Hi-Di formamide denatured for 2 minutes at 94°C and immediately run on an ABI 3100 sequencer. Note if the sequencing reaction signal is too high on the ABI 3100 samples should be serial diluted into a new 96-well plate - aiming for a RAW signal of 1000-2000.

2.9 Sequenom MassARRAY EpiTYPER

Sequenom's MassARRAY® EpiTYPER® is a bisulphite-treatment-based method for detection and quantitation of DNA methylation using base-specific cleavage and Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry (MALDI-TOF MS).

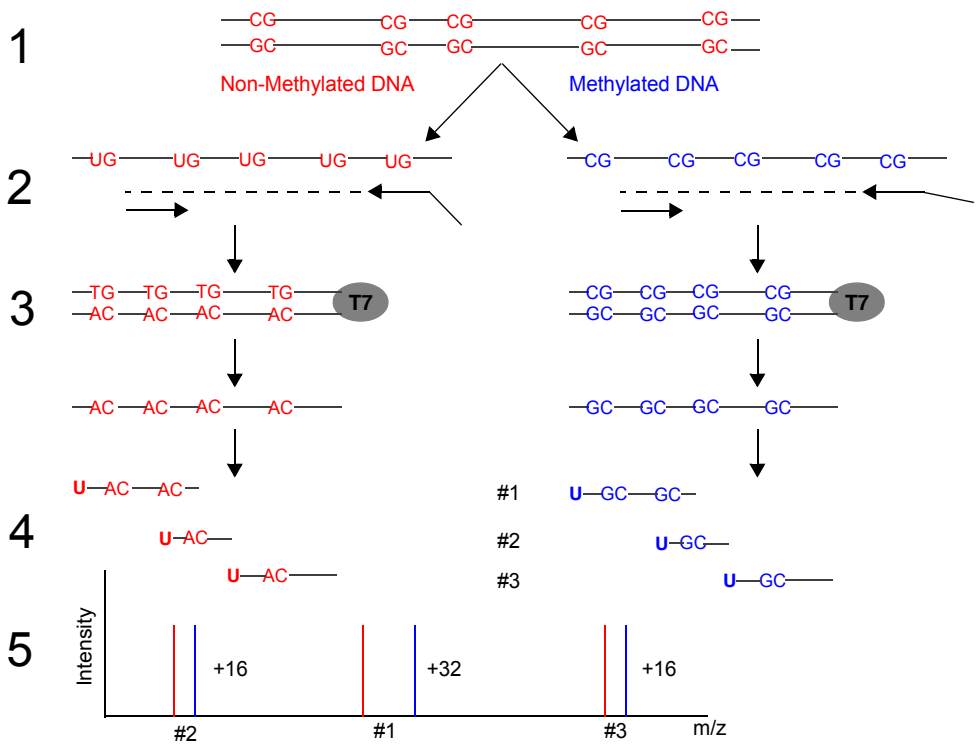


Figure 2.2 Overview of EpiTYPER procedure [1: Bisulphite treatment; 2: PCR and SAP treatment; 3: *in vitro* transcription; 4: RNase A uracil-specific cleavage; 5: MALDI-TOF MS. Adapted from EpiTYPER application guide, March 2011]

Bisulphite treatment produces methylation-dependent sequence variations of C to T in the PCR amplification products. Following a SAP clean up step of the PCR products, *in-vitro* RNA

transcription is performed on the reverse strand (where thymidine bases are replaced with uracil bases) with subsequent base-specific cleavage at the uracil bases by RNase A, yielding fragmented RNA molecules of the reverse transcription reaction. Essentially the C/T variations appear as G/A variations in the cleavage products generated from the reverse strand by base-specific cleavage. These G/A variations result in a mass difference of 16 Daltons per CpG site, which is detected by the MassARRAY system. In the mass spectrum, the relative amount of methylation can be calculated by comparing the signal intensity between the mass signals of methylated and non-methylated template.

Primers and assays were designed and validated as outlined in section 2.7. With the addition of a 10 base tag (5'-aggaagagag-3') on the forward primer and a T7-promoter tag (5'-cagtaatacgactcactatagggagaaggct-3') on the reverse primer (Figure 2.3).

Figure 2.3 Sequenom primer design with reverse strand transcription T7 tag
[Fwd: forward primer with green 10 base tag; Rev: reverse primer with blue T7-promoter tag]

1. After performing bisulphite PCR, check 2 µL of each sample including the controls on a 2% agarose gel, before proceeding with the following setup.
2. Transfer 5 µL of the samples into a 384-well plate, noting the sample transfer positions ready to dephosphorylate unincorporated dNTP's from the PCR using shrimp alkaline phosphatase (SAP).

3. Prepare a SAP PCR clean up mix as below and then split the mix into an 8 well strip tube.

Component	Volume for a single reaction	Volume for a 384-well plate (400 reactions)
RNase free water	1.70 µL	680 µL
Shrimp alkaline phosphatase (SAP)	0.30 µL	120 µL
Total volume	2.00 µL	800 µL

SAP clean up mix

4. Using the multi-channel pipette, add 2 µL of SAP mix to the 384-well plate, seal the plate and spin the contents down in a centrifuge.
5. Incubate as follows: 37^{o}C for 40 minutes followed by 85^{o}C for 5 minutes and cooled to 10^{o}C for 10 minutes.
6. Transfer 2 µL of SAP-PCR product into a new 384-well plate.
7. Prepare the transcription/RNase A mix as below for the T-specific cleavage reaction and split into an 8 well strip tube.

Component	Volume for a single reaction	Volume for a 384-well plate (400 reactions)
RNase free Water	3.21 µL	1284 µL
Polymerase Buffer [5X]	0.89 µL	356 µL
Cleavage Mix (T mix)	0.22 µL	88 µL
DTT [100 mM]	0.22 µL	88 µL
T7 RNA & DNA Polymerase [50 U/µL]	0.40 µL	160 µL
RNase A	0.06 µL	24 µL
Total volume	5.00 µL	2000 µL

8. Using a multi-channel pipette, add 5 µL of the mix to the 384-well plate containing the 2 µL of SAP-PCR product, seal the plate and spin the contents down in a centrifuge.
9. Incubate the sample plate at 37^{o}C for 3 hours.
10. Add 20 µL of double distilled water to each well and spinning down.

11. Using the 384-well Dimple plate 6 mg CLEAN resin aliquots into each dimple and leave to dry for 20 minutes.
12. Carefully tip sample plate upside down (surface tension should keep product in wells), place over dimple plate, making sure it is flush against the sides, turn both plates over so sample plate is on the bottom. Tap the dimple plate so resin enters each well of the sample plate.
13. Rotate the plate on the lowest setting for 10 minutes and then centrifuge for 10 minutes at 3800 rpm.
14. Processed samples are then transferred to the Sequenom SpectroCHIP using the MassARRAY nanodispensor.
15. Setup a sample sheet with assay and sample information using the EpiTYPER software along with any associated files required.
16. Using the EpiTYPER software on the instrument PC open up saved sample sheet and load onto the instrument software. Specify the MassCLEAVE parameter files and any reaction specifics.
17. Load the SpectroCHIP into the instrument and initialise. When ready select the ‘Run’ button.
18. Once the run is complete analyse the data in the EpiTYPER software using the T cleavage reaction setting.
19. Data once analysed are presented in the ‘EpiGrams’ tab - for a visual % methylation representation and ‘Methylation’ tab - tabulated % data. Depending on the desired analysis various images and measures can be extracted using the software and can be further analysed in the R package MassArray.

2.10 SERT promoter methylation analysis

Serotonin Transporter (SERT, SLC6A4, 5-HTT) gene promoter methylation analysis was carried out by Sequenom EpiTYPER. The following primers were designed and validated as outlined in section 2.7 and used to generate bisulphite PCR products for the Sequenom platform using the conditions below.

Primer	Sequence (5'-3')
SERT P1a Fwd	aggaagagagTATTGTTAGGTTTTAGGAAGAAAGAGAGAG
SERT P1a Rev	cagtaatacgactcactatagggagaaggctAACCCTCACATAATCTAATCTCTAAATAACC
SERT P1b Fwd	aggaagagagGGTTATTTAGAGATTAGATTATGTGAGGGTT
SERT P1b Rev	cagtaatacgactcactatagggagaaggctCCTACAACAATAAACAAAAAAACCCC

Sequences in lower-case = platform specific tags

Component	Final concentration
Qiagen PCR buffer	1X
MgCl	2mM
dNTP mix	100µM each
Primer F/R mix	0.05µM each
HotStar Taq	0.4U
DNA	20ng
PCR grade water	To a final volume of 10 µL

SERT P1a PCR mix

Component	Final concentration
Qiagen PCR buffer	1X
MgCl	2.5mM
dNTP mix	100µM each
Primer F/R mix	0.02µM each
HotStar Taq	0.4U
DNA	20ng
PCR grade water	To a final volume of 10 µL

SERT P1b PCR mix

Step	Temperature	Time
Initial	95°C	15 min
Denaturing	95°C	20 sec
Annealing	*63>57°C	1 min
Extension	72°C	1 min
Final	72°C	3 min
45 Cycles in a G-storm PCR machine		

*Cycling conditions for the annealing temperature involved a touchdown protocol, whereby an initial annealing temperature of 63°C was reduced by 1°C each cycle until 57°C, which was used thereafter.

2.11 Pyrosequencing using the PyroMark Q24

Bisulphite PCR primers and assays were designed and validated as outlined in section 2.7. Alternatively they were designed or checked using the PyroMark Assay Design Software. In addition a Biotin label is added to either the forward or reverse primer, this being dependent on the choice of strand the assay is designed on. The following diagram shows a schematic of the method where a third primer is used for the actual pyrosequencing part of the protocol.

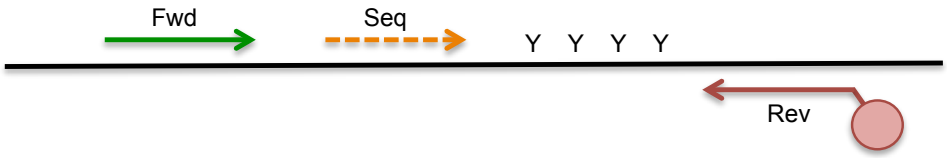


Figure 2.4 Illustration of a typical PyroMark Q24 CpG assay [Fwd: forward primer green; Rev: reverse primer red with biotin-tag; Seq: sequencing primer dashed orange; Y: CpG sites]

Generally a single 20 µL bisulphite PCR reaction or two separate 10 µL reactions per sample were setup and then pooled prior to running on the instrument as set out in the following protocol.

1. Set up a run file using PyroMark Q24 software specifying the 'method 003', choosing the assay with corresponding nucleotide dispensation order. Paste in the sample information from an Excel spreadsheet. Print the pre-run information and save the file on a USB memory stick.
2. Make up a mix of streptavidin beads and binding buffer as follows minus the PCR product.

Component	Volume for a single reaction	Volume for a run of 24 samples (25 reactions)
Streptavidin beads	2 µL	50 µL
Binding buffer	40 µL	1000 µL
PCR product	8-38 µL	200-950 µL
PCR grade water	To a final volume of 80 µL	To a final volume of 2000 µL

Pyrosequencing streptavidin bead mix

3. Aliquot the streptavidin bead/buffer mix to a three-column section cut off of a 96-well plate marking a corner of the plate for orientation.
4. Add between 8 µL and 38 µL of PCR product depending on PCR efficiency and assay and cover with a thick seal.
5. Place on the mixer, secure and mix at 1400 rpm for a minimum of 10-15 minutes.
6. Make up the sequencing primer mix as follows in a 1.5 mL tube.

Component	Volume for a single reaction	Volume for a run of 24 samples (25 reactions)
Sequencing primer [10 µM]	0.75 µL	19 µL
Annealing buffer	24.35 µL	609 µL
Total volume	25 µL	628 µL

7. Aliquot 25 µL of the primer mix to each well of a PyroMark Q24 plate and place in position on the Wash Station.
8. Set up the Wash station; add 50 mL of each Wash station reagent (70% ethanol, Denaturation solution, Wash buffer and high-purity water) as marked on the individual troughs.
9. Switch on vacuum pump, place filter probes on the PyroMark vacuum Prep Tool in high-purity water at the Wash station, run the water through to prepare the probes. Replenish the water.
10. PCR products are then taken off the shaker and immediately immobilised onto the PyroMark vacuum Prep Tool filter probes, purified, washed and denatured while attached to the PyroMark vacuum Prep Tool at the Wash station and finally the samples are released onto the prepared PyroMark Q24 plate.
11. Place PyroMark plate onto a heat block at 80 ^{o}C for 2 minutes.
12. Remove plate from heat block and place in the Pyrosequencer and wait for 5 minutes before starting the run.
13. Load the Pyrosequencer cartridge reagents as specified in pre-run information obtained from step 1. Place the cartridge in the instrument; attach USB memory stick and load the run file as instructed on-screen.
14. Once completed analyse the run using the PyroMark Q24 software in CpG mode. The quantification of CpG methylation and quality assessment are displayed above each CpG site in the pyrogram trace.

2.12 5-mC/5-hmC global methylation by ELISA

ELISA quantification of both global 5-methlycytosine (5mC) and global 5-hydroxymethycytosine (5hmC) was carried out using MethylFlash (colorimetric) kits from Epigentek. All DNA samples were quantified by Quant-iT PicoGreen fluorometry (see section 2.5) and equalized prior to testing in duplicate as set out in the kit protocol. In short, DNA is bound to strip wells that are specifically treated to have a high DNA affinity. The methylated fraction of DNA is detected using capture and detection antibodies and then quantified colorimetrically by reading the

absorbance in a microplate spectrophotometer. The amount of methylated DNA is proportional to the OD intensity measured.

2.13 Alu & LINE1 consensus CpG global methylation

Genome wide 5mC at CpG sites in consensus Alu and LINE1 repeat sequences was carried out by bisulphite PCR and Pyrosequencing (see section 2.11). DNA samples along with 0% and 100% methylated DNA controls (Qiagen; NEB) were bisulphite treated (see section 2.6.2) and setup using the following protocol.

1. PCR reactions are set up in a 20µL volume and cycled as below.

Primer	Sequence (5'-3')
Alu Fwd	Biotin-TTTTTATTAAAAATATAAAAATT
Alu Rev	CCCAAACTAAAATACAATAA
Alu sequencing	AATAACTAAAATTACAAAC
LINE1 Fwd	TTTTGAGTTAGGTGTGGGATATA
LINE1 Rev	Biotin-AAAATCAAAAAATTCCCTTTC
LINE1 sequencing	AGTTAGGTGTGGGATATAGT

Assay	Run file nucleotide dispensation order
Alu	G/AC/TG/AC/TG/ACCACCA
LINE1	GTCGTGTAGTCAGTCG

Component	Final concentration
Qiagen PCR buffer	1X
MgCl	3.5mM
dNTP mix	100µM each
Primer F/R mix	1µM each
HotStar Taq	0.8U
DNA	20ng
PCR grade water	To a final volume of 20 µL

Alu PCR mix

Component	Final concentration
Qiagen PCR buffer	1X
MgCl	1.5mM
dNTP mix	100µM each
Primer F/R mix	1µM each
HotStar Taq	0.8U
DNA	20ng
PCR grade water	To a final volume of 20 µL

LINE1 PCR mix

Step	Temperature	Time
Initial	95°C	15 min
Denaturing	95°C	30 sec
Annealing	50°C	30 sec
Extension	72°C	30 sec
Final	72°C	5 min
45 Cycles in a Dyad PCR machine		

LINE1 & Alu PCR cycling conditions

2. PCR products are then purified, washed and denatured using the PyroMark vacuum Prep Tool prior to running them on a PyroMark 24 (Qiagen) with 0.3µM of the assay specific sequencing primer (see section 2.11).
3. Pyrograms are checked visually for QC measures and then percentage methylation values exported using the PyroMark Q24 CpG assay analysis software. For both the Alu and LINE1 mean percentage for an individual sample is calculated using the three consensus CpG sites adjacent to the sequencing primer.

2.14 GLUMA (Glucosylation LUminometric Methylation Assay)

Global DNA methylation at -C*CGG- sites was assessed using a modified LUminometric Methylation Assay (LUMA) (Karimi et al. 2006). Essentially an additional MspI+EcoRI digestion is carried out with a pre-treatment with T4 Phage β-glucosyltransferase to allow for the discrimination of 5mC and 5hmC. Also we incorporated the measurement of known methylation content standards (Section 2.15.1) to generate a standard curve to quantify the amounts of the different cytosine species.

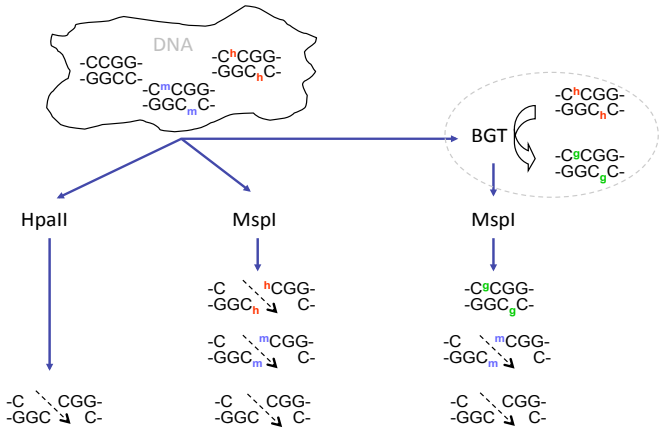


Figure 2.5 GLUMA sample preparation
[DNA is firstly glucosylated using T4 phage β-glucosyltransferase (BGT) then digested with MspI, alongside two further digestions of the same DNA by the HpaII & MspI isoschizomers, resulting in three different samples comprising digested: 5mC+C, C and 5hmC+5mC+C respectively, which are then detected and quantified using the pyrosequencer platform]

2.14.1 Glucosylation of 5hmC

The pre-treatment of DNA with T4 Phage β-glucosyltransferase was used to add a glucose moiety to the hydroxyl group of 5hmC. This modification protects the 5hmC within the -CCGG- context from digestion by the methylation insensitive MspI restriction endonuclease. Thus allowing the separation of 5hmC from 5mC based on the differential restriction enzyme digestion of -CCGG- sequences throughout the genome depending on the absence or presence of a modified cytosine at the internal position (Figure 2.5).

Glucosylation reactions are set up in a 384-well plate as detailed below and incubated for 2 hours at 37°C. All samples along with dsDNA oligo standards (see section 2.15.1) are setup in duplicate.

Component	Final concentration
NEB buffer 4	1X
UDP-Glucose	40µM
T4 Phage β-glucosyltransferase	3U
DNA	250ng
PCR grade water	To a final volume of 7.5 µL

Glucosyltransferase reaction setup

1. Following glucosylation, the reactions are ethanol precipitated. Add 2.5 µL of 125 mM EDTA to each well; tap the plate to ensure the EDTA reaches the bottom of the wells.
2. Add 25 µL 100% ethanol, seal the plate and mix by inverting the plate 4 times. Incubated at room temperature for 15 minutes prior to -20°C for 1 hour.
3. Then centrifuge at 1500 x *g* at 4°C for 45 minutes. Remove the plate seal and invert the plate onto a paper towel, centrifuge at 185 x *g* for 30 seconds.
4. Remove from the centrifuge and add 30 µL 70% ethanol and centrifuge at 1500 x *g* at 4°C for 15 minutes. Remove the ethanol by a short inverted spin for 1 minute as before.
5. The samples are briefly air dried for 5 minutes and immediately setup for MspI+EcoRI digestion as described below for the LUMA protocol.

2.14.2 LUMA

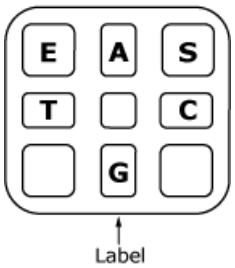
LUminometric Methylation Assay (LUMA) (Karimi et al. 2006), is an adaption of traditional isoschizomer endonuclease restriction digestion methods used to quantify the products generated from methylation sensitive (HpaII) and methylation insensitive (MspI) digestion using the pyrosequencer platform.

1. In brief, samples from the glucosylation step above are cleaved with MspI+EcoRI restriction enzymes, alongside these the same DNA samples (not treated for glucosylation) are also cleaved with MspI+EcoRI and HpaII+EcoRI in two further separate reactions as set out below and incubated for 16 hours at 37°C.

Component	Final concentration
Fermentas Tango buffer	1X
EcoRI	1.25U
MspI or HpaII	1.25U
DNA	250ng
PCR grade water	To a final volume of 7.5 µL

MspI+EcoRI & HpaII+EcoRI restriction digest setup

2. As with the glucosylation step DNA sample digestions are carried out in duplicate. The duplicates are pooled prior to processing in AQ mode on the PyroMark Q24 as outlined below.
3. Set up the LUMA assay run file using PyroMark Q24 software, with the ACTCGA dispensation order for the nucleotides and save file to a USB memory stick.
4. Add 20 µL of Annealing buffer to each well in a PyroMark Q24 plate and then add each of the three different digestion products for each sample duplicate into the three rows of the plate as follows: with the glucosylation & MspI+EcoRI digestion products added to the 1st row; MspI+EcoRI digestion only products added to the 2nd row; and the HpaII+EcoRI digestion only products added to the 3rd row.
5. Prepare the pyrosequencer cartridge reagents as specified below for two runs.



E = 180 µL pyrosequencing enzyme
S = 180 µL pyrosequencing substrate
A = 100 µL nucleotide A + 100 µL water
C = 100 µL nucleotide C + 100 µL nucleotide G
G = 200 µL dH_2O
T = 100 µL nucleotide T (red) + 100 µL water

6. Load the PyroMark Q24 plate and cartridge into the instrument; attach USB memory stick and load the run file.
7. Peak heights for each sample and its three digestions are calculated and exported in a tab-delimited format using the PyroMark Q24 AQ assay interface. The ratios used for quantitative analysis of global DNA methylation are calculated as described below.
8. Open the exported data in Excel. The peak heights for all six nucleotide dispensations (numbered as 1–6 in the order of dispensation) are specified. First, perform a quality check of your data. The first three peaks, following the substrate peak, should be substantially higher than the last three (CGA), which should be very small and

preferably non-existent. If these last peaks are high, especially the second C and A dispensations, this usually indicates fragmented DNA (poor samples, which should be removed from further analysis).

9. Calculate the average of peak 1 and peak 3 for each digest to get an average peak height for EcoRI restriction.
10. Next calculate the ratio of peak 2 to the averaged EcoRI peak, for HpaII and MspI digestions separately, to get normalized values for both enzymes across the three separate digests.
11. Calculate the ratio of normalized HpaII to normalised MspI from the un-glucosylated digest to get the HpaII/MspI ratio. This value indicates the degree of total methylation (5mC+5hmC).
12. Calculate the ratio of normalized HpaII to normalised MspI from the glucosylated digest to get the HpaII/glucosylation-MspI ratio. This value indicates the degree of 5mC methylation.
13. The observed ratios of methylation (5mC+5hmC and 5mC) for the dsDNA oligo standards are plotted against their expected ratios and a standard curve generated for these known standards. Using this slope re-calculate the portions of 5mC+5hmC and 5mC methylation for each of the test samples.
14. These re-calculated 5mC+5hmC and 5mC ratios are then converted to a percentage (%) methylation by subtracting them from 1 and then multiplying by 100.
15. Percentage of 5hmC methylation in each sample is then calculated by subtracting the 5mC % from the total 5mC+5hmC %.

2.15 Preparation of cytosine(C), 5-methylcytosine (mC) and 5-hydroxymethylcytosine (hmC) standards

For the measurement of global methylation, control standard curves were generated using serial dilutions of synthetic controls. Two types of synthetic controls were made-up in the lab, one consisting of complementary 55 base oligos where the central cytosine was either standard (C) or modified (mC or hmC); and one consisting of a 356bp PCR amplicon where all cytosines were either standard (C) or modified (mC or hmC).

2.15.1 Preparation of dsDNA oligo standards

Three pairs of custom complementary oligos as shown below, with and without the modifications of 5-methylcytosine and 5-hydroxymethylcytosine at the internal cytosine of the HpaII/MspI (CCGG) restriction site; plus an EcoRI restriction site (GAATTC) were purchased from Eurogentec.

Forward strand: 5'-TATACGAATTCAACAGCTATGACCATGCCGGAGCGG ATAACAATTTCACACAGGA-3'
Reverse strand: 5'-TCCTGTGTGAAATTGTTATCCGCTCCGGCATGGTCATAGCTGTTGAATTC GTATA-3'

Three dsDNA oligos (representing 100% normal cytosine, methylcytosine and and hydroxymethylcytosine at the -C*CGG- sequence) were generated by mixing 1:1 molar ratio (final concentration of 5µM) in a microcentrifuge tube with annealing buffer containing salt; 1X TE, 10 mM Tris, 1 mM EDTA, 50 mM NaCl (pH 7.5- 8.0) followed by 95°C for 5 minutes on a heat block and allowed to gradually cool to room temperature while leaving the tubes in the block over-night. These were then diluted to 0.1µM and mixed in varying proportions of the three cytosine species to be used as reference standards as shown in the table below.

Well	Total modified C %	mC %	hmC %	C %
1	0.0	0.0	0.0	100.0
2	23.5	23.2	0.3	76.5
3	30.0	29.5	0.5	70.0
4	50.0	48.9	1.1	50.0
5	60.0	57.9	2.1	40.0
6	70.0	65.8	4.3	30.0
7	83.0	74.5	8.5	17.0
8	100.0	83.0	17.0	0.0

2.15.2 Preparation of 356bp PCR standards

These consist of PCR fragments of 356bp designed around a cytosine rich region of the APC gene promoter using the primers below. Three types of cytosine species product were created by PCR, incorporating either standard cytosine, methylcytosine or hydroxymethylcytosine into the sequence using different dNTP mixes for each of the three cytosine species.

Primer	Sequence (5'-3')
APC promoter Fwd	AAGGAATTCGAGGAAGGTGAAGCACTCAGTT
APC promoter Rev	AGAGAATTCAGGGTGAGACATGGAGAGAAGA

1. Initially a primary PCR is carried out using genomic DNA as template with standard dNTP's. These products are then cleaned up using the QIAquick PCR purification kit (Qiagen) and quantified using a Nanodrop spectrophotometer.

Component	Final concentration
Qiagen PCR buffer	1X
DMSO	3%
dNTP mix	200µM each
Primer F/R mix	0.5µM each
HotStar Taq	0.8U
DNA	20ng
PCR grade water	To a final volume of 10 µL

Primary APC PCR

Step	Temperature	Time
Initial	95°C	15 min
Denaturing	94°C	30 sec
Annealing	60°C	30 sec
Extension	72°C	30 sec
Final	72°C	5 min
35 Cycles in a G-storm PCR machine		

PCR cycling conditions

2. Using the template generated from the primary PCR a secondary PCR is carried out with three different dNTP mixes differing in their cytosine species: a standard dNTP mix with dCTP (Roche); a 5-methylcytosine dNTP mix with d5mCTP (Zymo); a 5-hydroxymethylcytosine dNTP mix with d5hmCTP (Zymo). Where the dCTP and d5hmCTP dNTP mix PCR's requires 20ng of elute from primary whereas the d5mCTP dNTP mix PCR PCR requires 70ng.

Component	Final concentration
Qiagen PCR buffer	1X
dNTP mix	200µM each
Primer F/R mix	0.5µM each
HotStar Taq	0.8U
DNA	20/70ng
PCR grade water	To a final volume of 10 µL

Secondary APC PCR

3. Finally these three different cytosine species products are cleaned up using the QIAquick PCR purification kit (Qiagen); checked on a 2% agarose gel and quantified

using a Nanodrop spectrophotometer before being diluted to 15ng/µL. These purified dilutions are mixed in varying proportions of the three different cytosine species and subsequently used as reference standards.

2.16 QIAquick PCR purification

1. Add 5 volumes of Buffer PB to 1 volume of the PCR sample and mix.
2. If pH indicator has been added to Buffer PB, check that the colour of the mixture is yellow. If the colour of the mixture is orange or violet, add 10µl of 3M sodium acetate, pH 5.0 and mix. The colour of the mixture will turn to yellow.
3. Place a QIAquick spin column in a provided 2 mL collection tube.
4. To bind DNA, apply the sample to the QIAquick column and centrifuge at 13,000 rpm for 30-60 seconds.
5. Discard flow-through. Place the QIAquick column back into the same tube.
6. To wash, add 0.75 mL Buffer PE to the QIAquick column and centrifuge at 13,000 rpm for 30-60 seconds.
7. Discard flow-through and place the QIAquick column back in the same tube.
8. To completely remove any residual ethanol from the Buffer PE, carry out an additional centrifugation at 13,000 rpm for 30-60 seconds.
9. Place QIAquick column in a clean 1.5 mL tube.
10. Elute DNA, by add 50 µL Buffer EB (10mM Tris·Cl, pH 8.5) to the centre of the QIAquick membrane, let the column stand for 1 minute and then and centrifuge at 13,000 rpm for 1 minute.

2.17 DNA Degradase

A nuclease mix DNA Degradase Plus (Zymo Research) was used to degrade DNA to its individual nucleoside components. As the nucleosides lack the negatively charged phosphate group this facilitates their quantification via combined Liquid Chromatography and Mass Spectromatry (LC-MS).

Component	Final concentration
DNA degradase plus buffer	1X
DNA degradase plus	2.5U
DNA	330ng
PCR grade water	To a final volume of 25 µL

DNA Degradase Plus reaction setup

2.18 Fluidigm 48.48 Access Array PCR based bisulphite sequencing library

The 48.48 Access Array Integrated Fluidic Circuit (IFC) system was used to produce a barcoded sample bisulphite PCR sequencing library. The procedure follows a two-step approach. In the first step, 48 samples x 48 amplicons were amplified on a single 48.48 Access Array. Target regions are amplified with target-specific (TS) primer pairs that have been tagged with common sequence tags.

After harvesting the PCR products from the 48.48 Access Array, the second PCR step is carried out in a standard PCR plate to attach sample-specific barcodes and Illumina adaptor sequences. By incorporating sample-specific barcodes, all 2304 PCR products from a single 48.48 Access Array are unique and thus can be pooled for subsequent clean up, quantification and normalisation steps and ultimately allow for highly parallel bisulphite sequencing in a single lane of an Illumina HiSeq instrument. In total six 48.48 Access Arrays were run to generate 144 APP gene regions/amplicons for each of the 96 test samples using the Fluidigm system. Target-specific primer sequences for the 144 amplicons are listed in Appendix B.

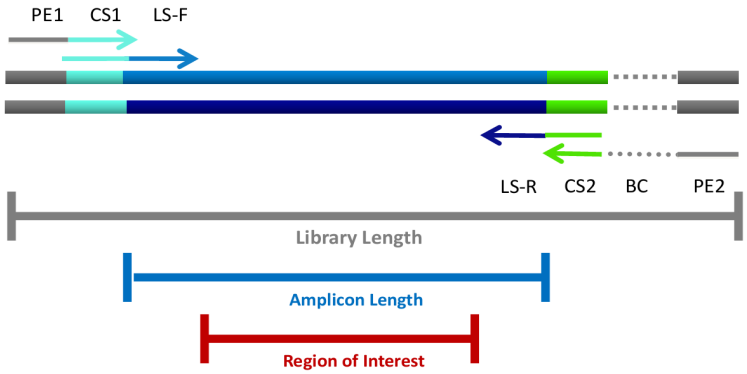


Figure 2.6 Illustration of the primers, barcode and PE adaptors incorporated in to PCR ROI [PE: paired end Illumina sequencing adaptor sequence; CS: common sequence between primary and secondary PCR primers; LS: locus specific primer sequence; BC: barcode sequence]

Protocol stage	Primer	Sequence (5'-3')
Primary PCR	CS1-TS-F	ACACTGACGACATGGTTCTACA–[TS-Fwd]
	CS2-TS-R	TACGGTAGCAGAGACTTGGTCT–[TS-Rev]
Secondary PCR	PE1-CS1	AATGATACGGCGACCACCGAGATCTACACTGACGACATGGTTCTACA
	PE2-BC-CS2	CAAGCAGAAGACGGCATACGAGA T-[BC]-TACGGTAGCAGAGACTTGGTCT

2.18.1 Designing targeted bisulphite sequencing primers

Given the large number of bisulphite PCR/sequencing primers that were needed to cover the desired region of interest, a semi-automated batch design approach was taken to design a set of 144 primer pairs.

1. Sizes of ROI sequences identified in the section above were manually checked and adjusted to 150bp where possible, while regions much larger than 150bp were broken up into a overlapping array of sequences.
2. These 150bp sequences +/- 50bp either side were extracted from the UCSC Human Mar. 2006 (NCBI36/hg18) assembly specifying, the reverse complement sequence (i.e. orientation of the APP gene in the minus strand) and known db135 SNP's to be identified with arrow brackets.
3. Sequences were formatted into an input file and Python scripting used to batch submit the individual sequences and to retrieve the primer design outputs from the web based bisulphite PCR primer design tool MethPrimer. The following MethPrimer parameters were chosen to standardise the design and subsequent amplification.

Parameter	Minimum	Optimum	Maximum
Product size:	130	150	175
Primer Tm:	54	56	58
Primer size:	20	25	33
Product CpGs:	1	-	-
Primer non-CpG C's:	3	-	-
Primer poly X:	-	-	5
Primer poly T:	-	-	8

2.18.2 Bisulphite conversion of 96 DNA samples

See section 2.6.2. A requirement of the standard 48.48 Access Array protocol is that the starting DNA template is 50 ng. This concentration is rather hard to achieve for bisulphite converted DNA and Nanodrop spectrophotometry only gives a rough quantification. However given this, the estimated concentration by Nanodrop measurement indicated that the samples ranged from 10-20 ng/µL so they were kept at room temperature for 30 minutes to allow for some evaporation and to concentrate a little. Following this aliquots of 2.3 µL of the eluted samples were made into columns 1-6 of a 96-well PCR plate and stored at -70^{o}C until further use.

2.18.3 48.48 Access Array primary target-specific bisulphite PCR

1. Thoroughly defrost a vial of the 1X Access Array Harvesting Solution and the 20X Access Array Loading Reagent to RT for an hour prior to use.
2. Prepare the 48.48 Access Array by injecting Control Line Fluid into each accumulator on the IFC.
3. Add 500 µL of 1X Access Array Harvest Solution into the H1-H4 wells on the IFC.
4. Load the IFC into the Pre-PCR IFC Controller AX located in the Pre-PCR lab, select Prime (151x) and Run Script to prime the IFC.
5. Once the script is complete, leave IFC in Controller AX until step 9 below.
6. Make up a 20X primer dilutions for each primer pair as outlined below. NOTE: The final Tagged TS forward and reverse primer concentrations are 1 µM in the 20X primer dilution. The final TS forward and reverse primer concentrations in the Access Array reaction chamber are 50 nM per primer.
7. Seal the 96 well plate(s) and vortex and spin down the 20X primer dilutions for 30 seconds.

Component	Final concentration
CS1-TS-Fwd	1 µM per primer
CS2-TS-Rev	1 µM per primer
Access Array Loading Reagent	1X
1X Te buffer	To a final volume of 20µL

20X primer dilution of Target-Specific primers

[N.B. 1X Te buffer has 1/10th the concentration of EDTA compared to standard 1X TE buffer]

8. Defrost all PCR components, vortex and spin down. Make up 65 reactions of the primary PCR mix as outlined below.

Component	Final concentration
Qiagen PCR buffer	1X
MgCl	5.5mM
dNTP mix	200µM each
Access Array Loading Reagent	1X
HotStar Taq	0.25U
PCR grade water	To a final volume of 3µL

Primary PCR mix

9. Vortex and spin down the mix. Aliquot 3 µL of the primary PCR mix into each of the 48 bisulphite converted DNA samples from step 5 above.
10. Seal the 96 well plate and vortex and spin down the PCR mix/template for 30 seconds.

11. Carefully pipette 4uL of the 20X primer dilutions in to the 48 primer inlets on the 48.48 Access Array IFC.
12. Carefully pipette 4uL of the PCR mix/template in to the 48 sample inlets on the 48.48 Access Array IFC.
13. Load the IFC into the Pre-PCR IFC Controller AX, select Load Mix (151x) and Run Script to load the reagents into array.
14. Once the script is complete, press Eject and remove the IFC.
15. Remove and discard the blue protective film from the bottom of the AA IFC.
16. Proceed to the thermal cycler and place the 48.48 Access Array IFC onto the cycler and start PCR by selecting the appropriate protocol.

Step	Temperature	Time	Cycles
Denature	50°C	2 min	1
	70 °C	20 min	1
	95 °C	15 min	1
Denature Annealing Extension	95°C 60 °C 72 °C	15 sec 30 sec 1 min	10
Denature Denature Annealing Extension	95°C 80 °C 60 °C 72 °C	15 sec 30 sec 30 sec 1 min	2
Denature Annealing Extension	95°C 60 °C 72 °C	15 sec 30 sec 1 min	8
Denature Denature Annealing Extension	95°C 80 °C 60 °C 72 °C	15 sec 30 sec 30 sec 1 min	2
Denature Annealing Extension	95°C 60 °C 72 °C	15 sec 30 sec 1 min	8
Denature Denature Annealing Extension	95°C 80 °C 60 °C 72 °C	15 sec 30 sec 30 sec 1 min	5

48.48 Access Array cycling conditions

17. After the PCR has finished, move the 48.48 Access Array IFC into the Post-PCR lab for harvesting.
18. Remove the remaining 1X Access Array Harvest Reagent from the H1-H4 wells.
19. Pipette 600 µL of fresh 1X Access Array Harvest Reagent into the H1-H4 wells.
20. Pipette 2 µL of 1X Access Array Harvest Reagent into each of the Sample Inlets on the IFC.

21. Load the IFC into the Post-PCR IFC Controller AX. Select Harvest (151x) and Run Script, once the script is complete, press Eject to remove the IFC.
22. To a labelled a 96-well plate carefully transfer 10 µL of harvested PCR products from each of the Sample Inlets into columns 1-6 of a 96-well PCR plate, using an 8-channel pipette. Store this at -20 ^{o}C until further use.

2.18.4 Fluidigm secondary barcode/adaptor PCR:

1. Defrost all PCR components, vortex and spin down. Make up the secondary PCR mix minus the Primers and the template, as outlined below.

Component	Final concentration
Qiagen PCR buffer	1X
MgCl	3.5mM
dNTP mix	200µM each
Primer BC_F/R mix	0.4µM each
HotStar Taq	0.25U
Primary PCR product	1 µL
PCR grade water	To a final volume of 5µL

Fluidigm barcode/adaptor Secondary PCR

2. Vortex and spin down the mix. Aliquot 3 µL of the secondary PCR mix into a 384-well plate to this pipette 1 µL of the sample specific barcode/adaptor primer from the Fluidigm Barcode Plate A1.
3. Finally add 1 µL of the harvested primary target-specific bisulphite PCR, seal the plate and cycle as follows:

Step	Temperature	Time
Initial	95°C	15 min
Denaturing	95°C	15 sec
Annealing	60°C	30 sec
Extension	72°C	1 min
Final	72°C	3 min
8 Cycles in a dyad PCR machine		

PCR cycling conditions

2.18.5 Post-PCR purification and quantification

The quality of the PCR products prepared on an Access Array IFC is critical for successful amplicon sequencing. Any contamination of primers or primer dimers in the PCR products will be directly reflected in the quality of sequencing reads. Therefore, the PCR products generated on a 48.48 Access Array IFC should to be qualified and purified before sequencing.

The PCR products generated on the 48.48 Access Array IFC are first analyzed using an Agilent 2100 Bioanalyzer to check the quality of the PCR products. Next, the PCR products from each 48.48 Access Array IFC are pooled together in equal volume to create a PCR product library. The PCR product library is then purified using AMPure XP beads. After purification, the PCR product libraries are accurately quantified by both the Agilent 2100 Bioanalyzer and Quant-iT PicoGreen fluorimetry systems and pooled before proceeding to cluster generation and sequencing.

2.18.6 Checking harvested PCR products on the Agilent 2100 Bioanalyzer

1. Run 1 µL of the 48 pooled PCR products from each sample on a Bioanalyzer DNA 1000 Chip (see section 2.4).
2. Check the results of the chip to determine if the PCR product in the DNA reaction has the expected size.
3. Ensure that amplicon sizes and distribution are within the expected range (+/- 5% for amplicons in the range of 200-400bp including tags). Primer dimer contamination in the PCR product pool (in the range of 50-130bp) should be less than 25% based on the Bioanalyzer quantification.

2.18.7 Purification of harvested PCR products by Ampure XP bead clean up

1. Remove Ampure XP beads from refrigerator and warm up at room temperature for 30 minutes.
2. Prepare a fresh 70% ethanol solution
3. For each 48.48 Access Array IFC run, pool 1 µL of each sample secondary PCR into a new 1.5 mL tube and mix gently by pipetting up and down.
4. Vortex Ampure XP beads for 10 seconds to resuspend. Bead solution should appear homogeneous and consistent in colour.
5. To a new 1.5 mL tube add equal volumes of the pooled sample PCR (36 µL) and Ampure XP beads (36 µL).

6. Mix gently by pipetting up and down for 20 seconds and incubate at room temperature for 10 minutes.
7. Place the 1.5 mL tube onto a magnetic separator and allow it to sit for 5 minutes.
8. Carefully pipette out the supernatant without disturbing the beads (remove as much liquid as possible).
9. Add 180 µL of 70% ethanol and vortex for 10 seconds.
10. Place the 1.5 mL tube onto a magnetic separator and allow it to set for 1 minute.
11. Carefully pipette out the supernatant without disturbing the beads.
12. Add 180 µL of 70% ethanol and vortex for 10 seconds.
13. Place the 1.5 mL tube onto a magnetic separator and allow it to set for 1 minute.
14. Carefully pipette out the supernatant without disturbing the beads.
15. Allow the beads to air dry for approximately 10 minutes by leaving the tube on the bench. Make sure the tube is completely dry before proceeding.
16. Add 40 µL of DNA suspension buffer to the tube and vortex for 5 seconds.
17. Place the 1.5 mL tube onto a magnetic separator and allow it to set for 2 minutes.
18. Carefully transfer the supernatant to a new 1.5 mL tube.

2.18.9 Quantification of cleaned products on the Agilent 2100 Bioanalyzer

1. Run 1 µL of the purified product library on a Bioanalyzer DNA 1000 Chip as described in section 2.4.
2. Define a region of interest in the electropherogram to determine the PCR product library concentration.
3. Select the 'Region Table' sub-tab on the bottom panel of the 'Electropherogram' tab. Right-click the electropherogram and select 'Add region'. Define the region to cover all of the PCR product library peaks. The 'Region Table' listed below the electropherogram will show the concentration of the region containing the PCR product library. Refer to the Agilent 2100 Bioanalyzer User Guide for additional information on Regions.

2.18.10 Quantification of cleaned products by PicoGreen fluorimetry

Using the estimated concentrations from above as a starting point quantitate the PCR product library by fluorimetry, using the Quant-iT PicoGreen dsDNA Assay Kit, as outlined in section 2.5.

2.18.11 Final PCR product library concentration

1. Given the PCR product library concentration (in ng/µL) from the methods above calculate the concentration in molecules/µL using the following equation:

$$\text{Molecules/µL} = \frac{(\text{PCR product conc.; ng/µL}) \times (6.022 \times 10^{23})}{(656.6 \times 10^{9}) \times (\text{average amplicon length; bp})}$$

   Where 6.022×10^{23} is Avogadro's number (molecules/mole), 656.6 is the average molecular weight of nucleotide pairs (g/mole), and average amplicon length (bp) is the average length of the 48 amplicons generated in the Access Array experiment.
2. Once each 48.48 Access Array IFC generated library has been purified and quantified, normalize the concentration of each library and pool libraries volumetrically by adding an equal volume of each library to a new 1.5 mL tube.
3. The pooled Access Array IFC library is now ready for sequencing.

2.19 Illumina 100bp PE sequencing

Illumina parallel sequencing was carried out at the Biomedical Research Centre - Genomics Core Facility, Guy's Hospital London. Using Qubit® Fluorometric Quantitiation the prepared library was re-quantified.

Subsequently 10 pM was loaded with 50% PhiX DNA spiked in and run on the sequencer with the following read configuration: 101bp Read 1 + 11bp Index read+ 101bp Read 2. Figure 2.7 shows the resulting sequencing coverage for an average 150bp amplicon by 100bp PE Illumina sequencing.

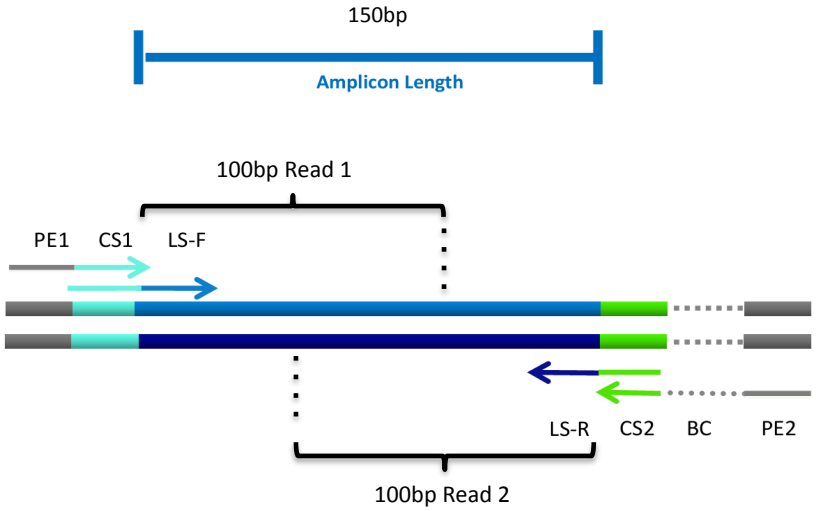
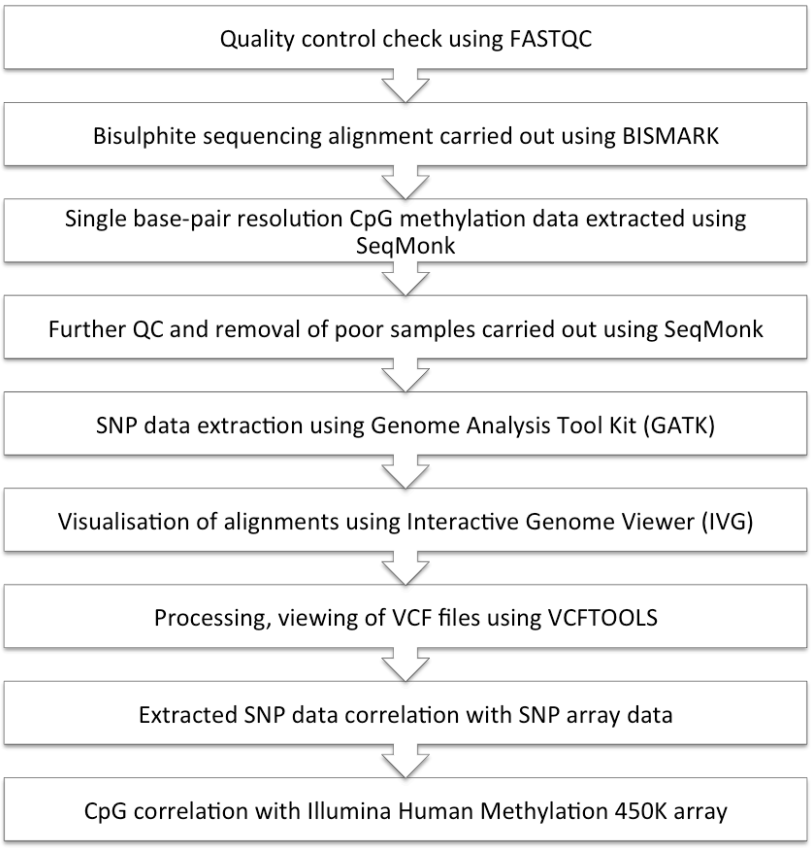


Figure 2.7 Illustration of reads 1 and 2 generated from 100bp PE Illumina sequencing [PE: paired end Illumina sequencing adaptor sequence; CS: common sequence between primary and secondary PCR primers; LS: locus specific primer sequence; BC: barcode sequence]

Once the sequencing was completed the single lane of data was extracted and converted using Casava v1.8.0. The barcodes are removed from the reads during demultiplexing and data files created for each of the de-barcoded 96 samples.

2.20 Parallel sequencing analysis pipeline



2.21 Statistical analysis

All statistical analyses were either performed using STATA statistical software version 12 (Stata Corporation, College Station, TX, USA) or R version 2.11.0 (R Foundation for Statistical Computing, Vienna, Austria). In general linear regression models were used to test the association of DNA methylation with genetic, demographic or clinical data. As a prelude to multivariate regression modeling, univariate regressions were carried out for individual independent variables (sex, age, ApoE genotype and experiment batch) with the dependent variable (DNA methylation) to ascertain if these variables needed to be added as a covariate to the final multivariate regression model; if a independent variable was significantly associated (p-value = <0.05) with the dependent variable it was included as a covariate in all further analyses. Unless otherwise stated all analyses were controlled for variables such as sex, age and ApoE genotype where applicable. Where data were available for more than one brain area mixed effects models were used to combine information from the different brain areas taking into account both between patient and within patient variation, using the data from the visual cortex brain area as baseline. The brain area data presented in the mixed model outputs merely summarise if the given brain area is different from the VC baseline area in that model. For all models, regression coefficients (β) represent the change in the dependent variable (such as DNA methylation) for each unit change in the predictor/independent variable. Where a large number of statistical tests were carried out on a given set of data, multiple testing corrections were applied using a Bonferroni correction, where the p-value of each test was multiplied by the total number of tests applied to the data and if the corrected p-value was still <0.05, the result was deemed to be significant.

MEASURES OF GLOBAL DNA METHYLATION IN POSTMORTEM BRAIN

3.1 Introduction

Global CpG methylation & Alzheimer's disease

DNA methylation can change over a person's lifespan, as part of normal aging, where the highest amounts of global methylation are observed in new-borns and gradually decreases through life (Heyn et al. 2012). While longitudinal studies in population based cohorts (Bollati et al. 2009; Bjornsson et al. 2008) have further demonstrated such changes in DNA methylation with time; twin studies highlight the separation of familial effects (i.e. genetic and common environment) and unique individual environment on the accumulation of such DNA methylation differences with age (Fraga et al. 2005; Wong et al. 2010; Talens et al. 2012), showing that these changes maybe attributable to the individual environment. These later life individual differences are likely to be due to unique environmental factors, such as diet and exposure to toxins or stochastic changes over time (Wallace et al. 2010; Bollati et al. 2007; Feil & Fraga 2012; Petronis 2010).

Furthermore, this phenomenon may be accelerated under certain conditions, as loss of global DNA methylation has been found in a variety of common human age-related diseases (Schumacher 2011). Recent observations of such global DNA methylation changes in AD post-mortem brain support the hypothesis that epigenetic mechanisms may have a role in the neurodegenerative process and alter AD risk (Mastroeni et al. 2009, 2010).

Masteroeni et al. (2011) note the epigenetic modifications that have been reported in AD, particularly with respect to DNA methylation (Mastroeni et al. 2009, 2010), typically resonate with similar trends in aging and may therefore help explain not only the pathologic complexity of AD, but also the particular salience of aging as an AD risk factor. Whether it is a shift in the normal DNA methylation profile at some point in time or a gradual "drift" over time (epigenetic drift), such changes driven by the environment may reflect more subtle changes in gene expression and thus help understand the aetiology and development of age related neurodegenerative diseases like Alzheimer's.

Other global methylation studies have suggested altered levels of repetitive element methylation as a blood marker for advanced age, namely a decrease in Alu DNA methylation with age (Bollati et al. 2009), while an increase in blood LINE1 DNA methylation was associated with AD (Bollati et al. 2011).

Transposed elements comprise 46% of the human genome, where retrotransposed Long Interspersed Nucleotide Element-1 (LINE1) represent the largest subgroup of repetitive elements. While most LINE1 elements are truncated copies with a defunct variable CpG rich 5' intrinsic promoter portion; 80-100 complete LINE1 elements exist (Kazazian 2004) that possess the potential to be active and it is commonly believed that they are kept suppressed by methylation (Rollins et al. 2006).

On the other hand Short INterspersed Elements (SINE) are approximately 100 to 400bp long and compared to other repetitive elements they are G+C and CpG rich throughout their entire sequence. In most species the G+C and CpG content of SINE elements is not enough to classify them as CGIs (Rollins et al. 2006). However a subclass of SINE sequences, Alu elements seem to be an exception in that they often overlap with CpG rich sequences many of which are CGIs, or at least exhibit sequence characteristics similar to CGIs (Paulsen et al. 2008). As revealed by the observation that, active domains of the genome show a marked enrichment of TSSs (Transcription Start Sites), CGIs and Alu sequences; whereas repressed domains of the genome are enriched with other classes of repeat sequences such as, LINE1 and LTRs (Long Terminal Repeats) (The ENCODE Project, 2007).

Additionally only a half of all CGIs distributed in the human genome are at gene promoter regions, the other half are dispersed within both intragenic and intergenic sequences (Illingworth et al. 2010) and though the role of non-promoter CGIs is yet to be fully understood, 42% of these CGIs show evidence for promoter activity (Faulkner et al. 2009; Illingworth et al. 2010). Further methylation mapping studies using human brain have shown that, these intragenic CGIs are significantly over-represented as tissue-specific differentially methylated regions (TS-DMRs) (Davies et al. 2010); and that the majority of methylated CGIs are in intragenic and intergenic regions associated with methylation-regulated alternative transcripts that are expressed in a tissue and cell type-specific manner (Maunakea et al. 2010). An interesting possibility is that repetitive sequences among such non-promoter CGIs may have some regulatory functions (Kim et al. 2009). Indeed Faulkner et al. (2009) have identified a large number of candidate regulatory regions that derive from retrotransposons, while retrotransposons located in the 5' UTR of protein-coding loci frequently function as alternative

promoters and/or express noncoding RNAs in a tissue specific manner, whereas those located in the 3' UTR of genes (~25% of RefSeq genes) can affect expression of those transcripts.

Overall Alu and LINE1 retrotransposon CpGs are largely methylated in normal somatic tissue. Demethylation of these elements is thought to contribute to the global demethylation associated with aging (Fraga 2009; Bollati et al. 2009) and given that together they account for around ~30% of the CpGs in the human genome they are often used as a surrogate measure of global DNA methylation levels. For these reasons and the possibility that these interspersed sequences in their own right play a role in the epigenome architecture, Alu and LINE1 consensus sequence CpGs were investigated in this chapter to identify possible differences in global DNA methylation levels with aging and AD. Further to this it would be interesting to see if the reported changes in methylation at these sequences (Bollati et al. 2009, 2011) for blood samples from AD patients are also present in post-mortem brain.

5-hydroxymethylcytosine

More recently the re-discovered modification of 5-hydroxymethyl-cytosine (5hmC) - a hydroxylated and methylated form of cytosine - has created a huge amount of interest in both developmental biology and neurobiology, as substantial amounts have been detected in both ES cells and the brain (up to ~40% of modified cytosines in brain) compared to other tissues (Kriaucionis & Heintz 2009). Across adult tissues, though 5mC levels are fairly consistent (~5% of all cytosines), 5hmC levels vary between <0.1% and ~0.7% and are highest in tissues from the brain and is particularly abundant in regions involved in higher cognitive functions, such as the cortex and hippocampus (Kriaucionis & Heintz 2009; Globisch 2010; Jin et al. 2011).

Kraus et al. (2012) has shown that 5hmC values in human cortical brain are highest in grey matter areas (mean 1.17%) compared to white matter (0.696%). Furthermore, looking at the hippocampal formation, a region of the brain affected in the early stages of AD (Hampel et al. 2008), immunolocalization of hippocampal 5hmC in mouse, has revealed this marker seems not to be uniformly distributed but prominently present in the dentate gyrus and CA1 cells (Chen et al. 2012). The genomic distribution of 5hmC in mouse ES cells shows that 5hmC is relatively enriched in euchromatic parts of the genome, including CGIs and promoters, and its presence in promoters and exons is associated with increased levels of transcription (Ficz et al. 2011).

Moreover in contrast to the low levels of 5mC in the CGI/promoter region of active genes, mapping of 5hmC in DNA of human frontal lobe revealed that 5hmC is selectively present in

active gene promoters and intragenic regions with significant enrichment at genes involved in neural development and patterning (Jin et al. 2011). Likewise observations of a trend where by 5hmC in the gene body increases with increased transcription of the corresponding gene whereas 5mC decreases was associated with synaptic plasticity genes (Khare et al. 2012). More specifically 5hmC densities are lower in alternatively spliced exons relative to constitutive exons, this being brain specific as compared to non-neuronal tissues which were noted to use 5mC rather than 5hmC for this proposed splicing related function (Khare et al. 2012).

With regards to aging some studies have shown an age related increase in global 5hmC in mouse brain (Münzel et al. 2010; Song et al. 2010; Chen et al. 2012) where this increase is not associated with DNA damage or oxidative stress, as previously suggested (Zawia et al. 2009). Age related increase of 5hmC in mouse cerebellum also revealed this increase was particularly pronounced in genes and pathways related to neurodegeneration (Song et al. 2010). Together these findings indicate that the aging process and possibly AD is associated with changes in 5hmC.

Though the exact role of 5hmC is unknown, it may be a modification in its own right attracting a unique panel of chromatin or transcriptional modifiers and thus adding another layer of complexity to the intricate network of epigenetic regulators (Branco et al. 2012). 5hmC reportedly inhibit the binding of various methyl-CpG-binding domain proteins such as MBD1, MBD2, MBD4 and MeCP2 (methyl-CpG-binding protein 2), thereby counteracting the role of certain 5mC-targeted transcriptional repressors (Jin et al. 2010).

Alternatively the transcriptional activity could be affected by proteins that selectively recognise 5hmC. Using a mouse knockout model of Rett syndrome, a neurological disorder caused by a mutation in the MeCP2 gene; Szulwach et al. (2011) demonstrated that 5hmC was inversely correlated with levels of MeCP2 and observed enrichment of 5hmC at LTR, LINE and SINE sequences, indicating that MeCP2 may protect conversion of 5mC to 5hmC. With the association of 5hmC with more highly expressed genes, these data hint at a mechanism by which expression of retrotransposon transcripts may be activated in the absence of MeCP2, suggesting that 5hmC is important in human neurological disorders (Szulwach et al. 2011).

Some evidence does show 5hmC to play a role in DNA demethylation, whereby 5hmC represents an intermediary form within the active demethylation pathway (Figure 3.1). Current models suggest either a passive or active demethylation. In proliferating tissues, DNMT1 has been found to poorly recognize 5hmC and therefore cannot maintain symmetric DNA

methylation patterns (Valinluck & Sowers 2007). Such that, a lack of replication of 5hmC by DNMT1 acts as a passive demethylation process (Guo et al. 2011; Ito et al. 2010). In the active TET-initiated demethylation model, the formation of 5hmC is catalyzed by the ten-eleven translocation (TET1-3) family of enzymes (Nabel & Kohli 2011). These enzymes have been linked to a number of different roles in embryogenesis as well as the elusive active DNA demethylation process. In relation to AD previous findings have suggested a nominally significant association between a TET1 polymorphism (rs5030882) and AD (Morgan et al. 2008).

The reason for the high levels of 5hmC in neuronal tissue is currently unknown. Perhaps it is the lack of cell division that prevents dilution of 5hmC by DNA replication, and 5hmC can therefore accumulate over time. It has also been suggested that the formation of 5hmC is a preventative measure against de novo methylation errors in neuronal cell types (Valinluck & Sowers 2007; Guo et al. 2011).

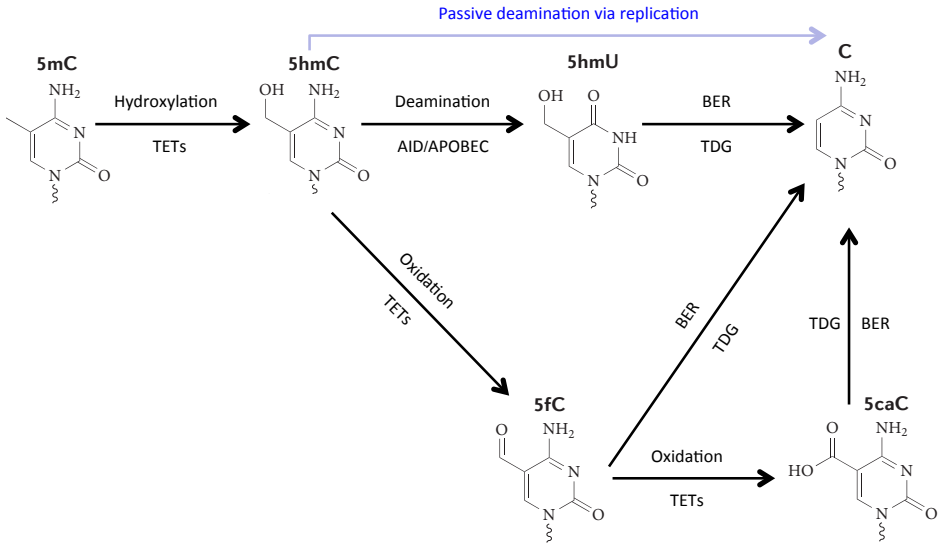


Figure 3.1 Proposed pathways for DNA demethylation
[The active demethylation pathway is proposed to involve hydroxylation of 5mC to 5hmC by TET1, 5hmC can be further oxidized by TET proteins to produce 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC). Alternatively, 5hmC may be further deaminated to become 5hmU by AID/APOBECC deaminases. 5hmU, 5fC, and 5caC can be excised from DNA by glycosylases such as TDG (thymine-DNA glycosylase). This base excision repair (BER) releases and replaces the entire modified oxidized base with unmodified C. A putative decarboxylase may also directly convert 5caC to C; Adapted from Tsukada 2012]

5mC and 5hmC detection

A number of different methods to measure global DNA methylation have become available in recent years and range from those that are direct measures of the total DNA methylation content, such as Enzyme-Linked ImmunoSorbent Assay (ELISA) based assays to those that are

indirect using bisulphite PCR and quantitative analysis by pyrosequencing at a given CpG dinucleotide sequence context (Lumey et al. 2012; Rath & Kanungo 1989; Wu et al. 2011; Gao et al. 2011). Although some antibody and enzyme based methods are still relatively expensive when planning to test a large number of samples, while bisulphite PCR based methods provide higher throughput, the major difference in traditional methods applied prior to the recent interest in 5hmC, is that 5hmC may go undetected by 5mC antibody methods, leading to under reporting of DNA methylation. Equally bisulphite based methods do not discriminate 5mC and 5hmC (Jin et al. 2010; Nestor et al. 2010).

The majority of current techniques for analysis of DNA methylation rely on the treatment of DNA with sodium bisulphite, where unmodified cytosine is converted to uracil. During subsequent PCR the unmodified uracil's are converted to thymine, where both 5mC and 5hmC will be read as cytosine. As a result the relative contributions of 5mC and 5hmC in the final signal is unknown, only giving a total of 5mC+5hmC. Furthermore it is reported that the bisulphite adduct of 5hmC, cytosine 5-methylenesulfonate (CMS) resulting from bisulphite treatment, stalls Taq polymerase during elongation, especially when two CMS bases are together (Huang et al. 2010) and this may even lead to biases in the PCR product. Therefore DNA regions containing tandem 5hmCs might be underrepresented through inefficient PCR following bisulphite treatment.

Accurately discrimination of 5mC and 5hmC content, can been achieved using chromatography based methods such as Thin Layer Chromatography (TLC), the method used in the original studies that detected the presence of 5hmC in brain (Kriaucionis & Heintz 2009; Tahiliani et al.2009) and Liquid Chromatography and Mass Spectrometry (LC-MS), which is regarded as the gold standard for global quantification of 5hmC (Globisch et al. 2010; Kraus et al. 2012; Le et al. 2011). Whereby exact quantification of hydrolysed nucleotides is achieved by the inclusion of isotopically labelled standards or simultaneously running a set of standards. However these methods require specialised equipment and expertise.

Another approach that can distinguish 5mC from 5hmC is by enrichment of DNA using antibodies that specifically recognize either 5mC or 5hmC. For example enrichment with antibodies has been used widely in combination with arrays and parallel sequencing to allow genome-wide mapping of either/both 5mC and 5hmC (Nestor et al. 2011; Wu et al. 2011). Discrimination between 5mC and 5hmC may also be accomplished by using methyl-CpG-binding proteins (MBD1, MBD2 and MBD4) that recognize 5mC at CpG dinucleotides but not 5hmC. However, these methods are highly CpG density-dependent, challenging in terms of sensitivity

and mapping accuracy, while they are more suited to examining the distribution of the relative 5mC or 5hmC content over the genome as opposed to a global quantitative measure of levels (Jin 2010). The other antibody-based method more suited to measurement of global DNA levels is ELISA. This was one of the first commercially available assays allowing differentiation of 5hmC and 5mC in separate kits (Epigentek and Zymo Research).

Some novel methods for the detection of 5hmC have been developed based on the selective glucosylation of 5hmC by the T4 phage β-glucosyltransferase (BGT) (Kinney et al. 2011). This enzyme selectively adds a glucose moiety to 5hmC, which is then either processed through a series of chemical reactions to enable the addition of biotin molecules to 5hmC or pulled down with and other protein, J-binding protein 1 (JBP1) that has a strong affinity for glucose-5hmC (Robertson et al. 2011). Much the same biotin addition to 5hmC allows the isolation of DNA containing 5hmC, which is then subjected to sequencing (Song et al. 2011b). Further modifications of this BGT reaction have used oxidation of the glucosylated 5mhC or differential restriction digestion based microarray mapping of 5hmC (Pastor et al. 2011; Khare et al. 2012). The use of BGT to differentiate 5mC and 5hmC is seemingly accurate and perhaps more widely accessible than LC-MS. While commercial kits (from NEB and Zymo Research) employing the use of BGT coupled with restriction enzymes, allow 5mC/5hmC quantification of specific genomic loci, they are further restricted to sites of enzymatic cleavage within a small region of interest.

Whereas methods using restriction enzymes for the detection of global DNA methylation, such as LUMA (Karimi et al. 2006) in their native form are also not specific for either 5mC or 5hmC, they can be adapted to be so. Such approaches based on the use of isoschizomer pairs of restriction endonucleases with different sensitivities to methylated cytosines, such as HapII/MspI at -C*CGG- and Psp6I/AjnI at -C*CWGG- can not distinguish the two forms of 5mC and 5hmC directly. However, it is possible that some enzymes are sensitive to selective modification of 5hmC, following glucosylation. For instance MspI at -C*CGG- was demonstrated to be resistant to digestion following the addition of a glucose moiety to 5hmC using BGT, whereas a bulkier group (glucose+biotin) was needed to block the enzyme TaqαI at -T*CGAA- (Song et al. 2011a). This was the approach taken in this chapter for the development of a modified assay based on the standard LUMA assay and pre-treating with BGT prior to MspI digestion.

Even though estimates of 5hmC represent a small fraction of the total cytosine content in most cells, this could still have a significant impact, particularly if 5hmC has important roles distinct

from 5mC, most notably in the brain as noted above. For this reason it is essential to distinguish 5hmC from 5mC. Although there are a number of techniques to investigate this, each method has its benefits and provides different insight into the global methylation content. While much information on the methylation status of specific loci or globally represented sequences has been generated to date particularly in the areas of cancer research and aging very little of this research has been carried out in the context of neuropathology and AD. Given that these global changes occur with aging the most salient risk factor for AD I chose to explore this in a case control manner in a cohort of post-mortem brains covering both a AD pathology brain area and non-AD pathology affected areas.

3.2 Aims

I. To measure DNA methylation at repetitive elements (Alu and LINE1) by pyrosequencing in post-mortem brain and to assess whether any differences observed are in line with those reported in blood samples (Bollati et al. 2011).

II. To measure levels of both 5mC and 5hmC global DNA methylation using two different methods. Firstly using commercial ELISA based whole genome methylation kits and secondly by a proposed modification of the LUMA assay at -CCGG- sequences across the genome.

III. To develop and asses the use of a modified LUMA assay to measure both 5mC and 5hmC global DNA methylation in a single assay.

IV. To ascertain if there are measurable differences in both the 5mC and 5hmC content of an AD pathology affected brain area and non-AD pathology affected areas across cases and controls.

3.3 Samples

All samples were taken from a well characterised cohort of brains donated to Medical Research Council London Neurodegenerative Diseases Brain Bank. These consisted of both AD brains and cognitively normal controls (N=66). AD brain samples were matched with normal controls on sex and age where possible, though this was limited by the small number of normal control brains available for analysis. The mean age at death was 86.3 years (Range: 69–89) for AD cases and 77.1 years (Range: 40–92) for controls. Twenty-seven patients were female (69%) and 12 males (31%), while in the control group, 15 subjects were females (56%) and 12 subjects were

males (44%). Frozen tissue of three different brain areas was collected, the Superior Temporal Gyrus (STG) as an AD pathology affected tissue, Visual Cortex (VC) as a non-AD pathology affected tissue and Brodmann Area 9 (BA9) as an area associated with behaviour. The choice for the STG and VC, would allow for comparison between AD pathology and non-AD pathology areas. DNA was extracted using a standard phenol-chloroform extraction (section 2.1). Though the BA9 region was chosen for its association with the behavioural phenotype in AD that may relate to global methylation changes, the number of samples tested by each method from this brain area are smaller than that those for the STG and VC. Samples from each brain area that were successfully assayed for each of the three main assay types are present in Table 3.1. A full list of the samples used in this chapter along with demographic details is listed in Appendix A.

Brain Area	Disease status	Sex	Alu/LINE1	ELISA	GLUMA
Visual cortex	AD	Female	26	15	24
		Male	10	5	9
	Ctrl	Female	11	8	14
		Male	8	8	12
Total			**55**	**36**	**59**
Superior Temporal Gyrus	AD	Female	25	15	24
		Male	10	5	10
	Ctrl	Female	12	8	14
		Male	9	8	12
Total			**56**	**36**	**60**
Brodmann Area 9	AD	Female	12	8	11
		Male	10	2	11
	Ctrl	Female	10	4	9
		Male	8	4	8
Total			**40**	**18**	**39**

Table 3.1 Numbers of samples successfully tested for global methylation by each method [AD: Alzheimer's disease; Ctrl: normal control]

3.4 Statistical analysis

In general linear regression models were used to test the association of global DNA methylation with, demographic data and AD status or Braak staging. Univariate regressions were carried out for individual independent variables (sex, age, ApoE genotype and experiment batch) with the dependent variable (DNA methylation) to ascertain if these variables needed to be added as a covariate to the subsequent multivariate regression models; if a independent variable was significantly associated (p-value = <0.05) with the dependent variable it was included as a covariate in all further analyses. In the first instance the separate relationships of age and sex alone on global DNA methylation levels were tested using all the combined data from each measure of global methylation in univariate linear regression, accounting for batch effects only. Secondly for each measure of methylation, data from individual brain areas in turn were sought

for an association with AD status or Braak staging by linear regression, controlling for sex, age, ApoE genotype and batch effect if necessary. Lastly mixed effects models were used to combine data from the AD pathology affected brain area (STG) and the non-AD pathology affected areas (VC and BA9), taking into account both between patient and within patient variation. Using the data from the visual cortex as baseline, differences in DNA methylation levels between the baseline and other brain areas were tested along with any association with AD status or Braak staging, again controlling for age, sex, ApoE genotype and batch effect where necessary.

3.5 Global methylation content assayed

First consensus sequence CpG methylation (5mC+5hmC) at both Alu and LINE1 sequences throughout the genome was measured via bisulphite PCR and pyrosequencing. Both assays quantified three CpG dinucleotides within a consensus sequence at the 5' region of these abundant sequences throughout the genome. Giving a snapshot of global methylation levels (Table 3.2), comprising 26% (Alu) and 11.8% (LINE1) of CpG within the genome. Note that both forms of methylation 5mC+5hmC contribute to the total level measured by these assays, as the bisulphite PCR/pyrosequencing method used does not distinguish between the two forms of methylation. Example pyrogram results are shown in Figure 3.2.

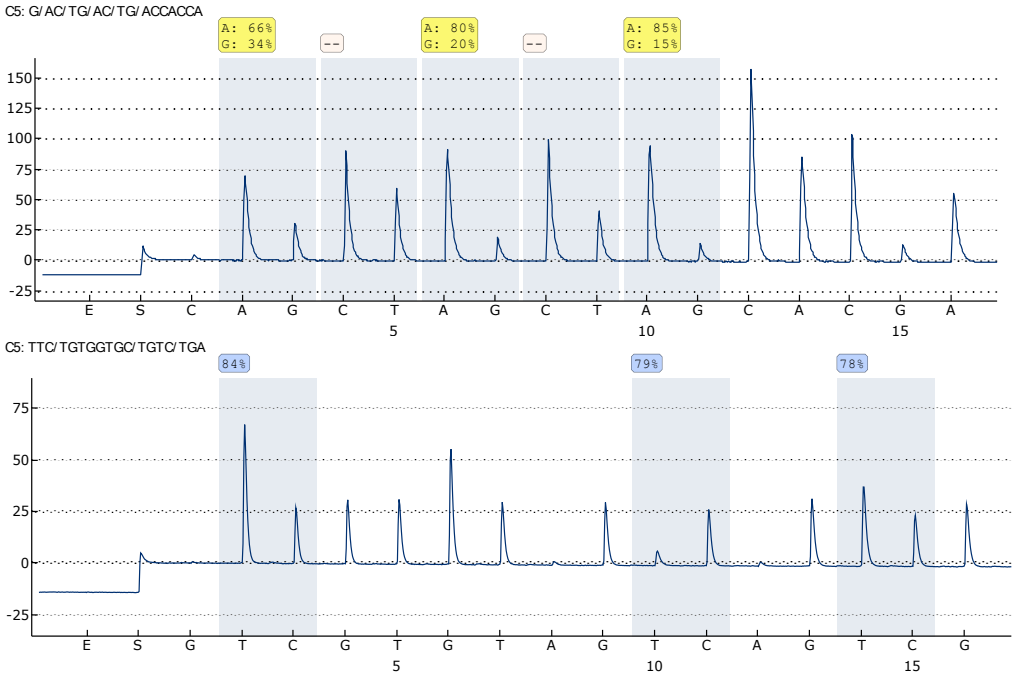


Figure 3.2 Example of Alu and LINE1 pyrogram results. [upper pyrogram of Alu result from a single sample, with yellow boxes showing quantification of the three CpGs dinucleotides tested; lower pyrogram LINE1 result from a single sample, with blue boxes showing quantification of the three CpGs dinucleotides tested]

Secondly quantification of total global 5-methlycytosine (5mC) and global 5-hydroxymethycytosine (5hmC) was carried out using MethylFlash (colorimetric) kits from Epigentek. The assays were expected to detect the total fraction of 5mC and 5hmC within the genome (Table 3.2) as they are based on specific 5mC or 5hmC antibody capture. Using these commercially available kits it was anticipated more accurate results with the added advantage of distinct detection of each type methylation would be ascertained.

Finally a modified LUMA assay with incorporation of a glucosylation step, referred to as GLUMA from here on, was developed to enable the discrimination 5mC and 5hmC, based on the measurement of total methylation as 5mC+5hmC and 5mC, followed by subtraction of 5mC levels from 5mC+5hmC to quantify 5hmC content. While this method does not measure total genome wide methylation like the ELISA kits, -C**C**GG- sequences are estimated to comprise ~25% of total CpGs within the human genome (Table 3.2) and cover ~90% of CGIs making them a good surrogate measure (Singer et al. 2010). Also unlike the ELISA kits this method would allow for the simultaneous measurement of both 5mC and 5hmC content in one assay. Example pyrogram results and standard curves generated from synthetic oligos to quantify methylation levels are shown in Figures 3.3 and 3.4. For a more detailed protocol of these three methods outlined above, see section 2.14 of Chapter 2.

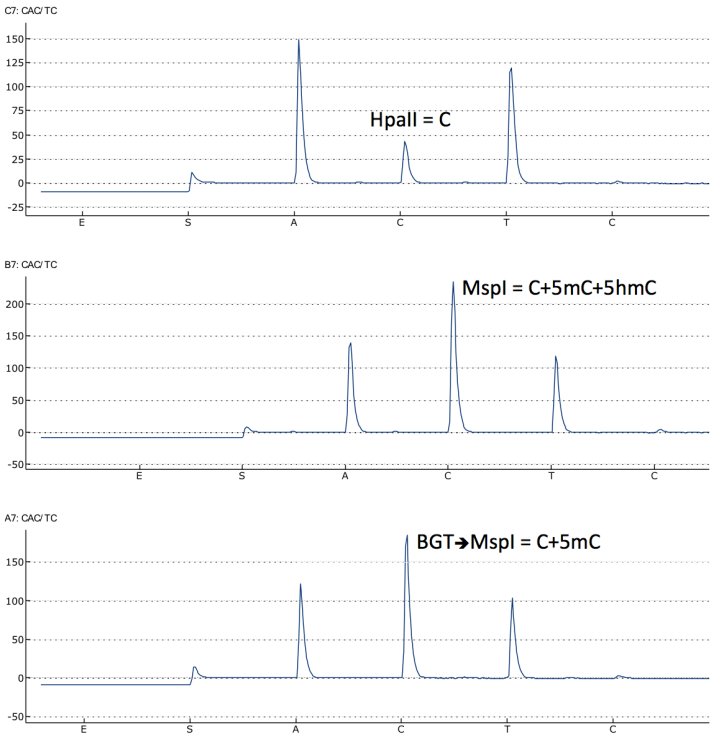


Figure 3.3 Example of GLUMA pyrogram results
[middle C peak in each pyrogram represents the signal of digestion products from the three respective digests for a single DNA sample; outer A and T peaks represent the internal control EcoRI digestion peaks]

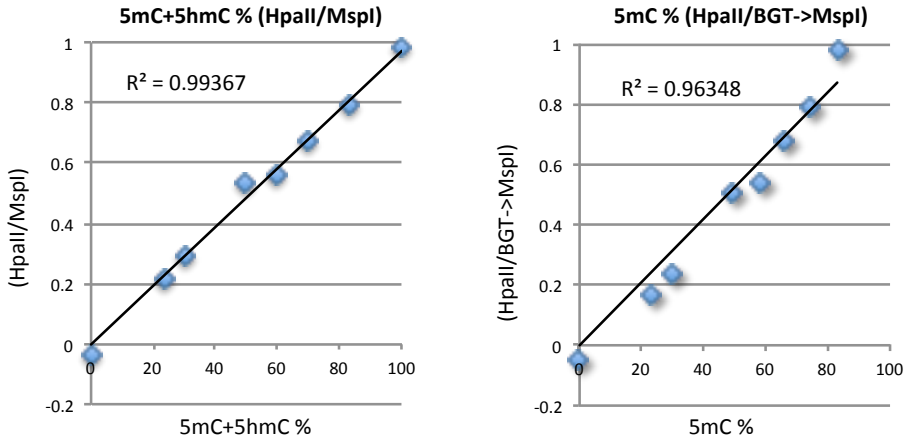


Figure 3.4 Example standard curves for total 5mC+5hmC and 5mC generated from a series of synthetic oligo dilutions of known methylation content

Table 3.2, gives a summary of the genomic methylation content assayed by each of the methods used. As each assay detects a different sequence context, the resulting measurements relate to the given percentage of CpGs within the genome and their corresponding portion of the total cytosine pool in the genome.

Method	Sequence context	% of CpGs in the genome assayed	Corresponding % of cytosines in the genome	5mC/5hmC discrimination
Alu	**C**A/G**C**A/G**C**A/GCCACCAC	26*	1.23	No
LINE1	**C**GTGGTG**C**GT**C**G	11.8*	0.56	No
ELISA 5mC	5m**C**	100	4.74	Yes
ELISA 5hmC	5hm**C**	100	4.74	Yes
GLUMA	-C**C**GG-	25[#]	1.18	Yes

Table 3.2 Summary of methylation content assayed by the different methods
[*: % based on figures from Rollins et al. 2006; [#]: % based figure from Mastroeni et al. 2011]

3.6 Results

3.6.1 Global methylation content assayed by method

Average levels of global methylation detected in the brain by each method are summarized in Table 3.3, highlighting the range of mean level values across the different assays. A mean methylation level in the range of 50-75% includes those for LINE1, GLUMA 5mC and GLUMA 5mC+5hmC, while all other measures are in the range of 0-25%, with the 5hmC levels lowest down to around 1%.

Methylation Measure	Mean (%)	Std. Err.	95% CI	
Alu	25.88	0.23	25.42	26.34
LINE1	74.22	0.29	73.65	74.79
ELISA mC	15.94	1.52	12.90	18.97
ELISA hmC	1.16	0.11	0.94	1.37
GLUMA mC+hmC	64.12	0.79	62.56	65.69
GLUMA mC	50.21	1.17	47.88	52.55
GLUMA hmC	13.91	0.73	12.46	15.35

Table 3.3 Summary of average methylation levels in brain measured by the different methods

As summarized in Figure 3.5, the global methylation content assayed in all the brain samples by the different methods correlate with a linear trend, with the exceptions of 5hmC as measured by the GLUMA assay against all other measures and Alu methylation against total methylation of GLUMA 5mC+5hmC, which show an inverse relationship. This weak inverse relationship of 5hmC against all other measures is what might be expected given the proposed role of 5hmC in de-methylation of 5mC. Results from the Alu assay were modestly correlated with both the measure of ELISA 5hmC and GLUMA 5hmC (r = 0.43, p =<0.001; r = -0.25, p =0.002 respectively). Methylation at Alu and LINE1 sequences correlate well (r = 0.32, p =<0.001) with one another and also show better correlation with those as measured by the ELISA assays (range 0.15-0.43) compared to GLUMA assay (range -0.063-0.15). Though the 5mC methylation levels measured by ELISA correlate with those obtained by the GLUMA assay (r = 0.33, p =0.002), the levels of 5hmC detected by both these methods bear little or no relationship (r = -0.13, p =0.23). The strongest correlations are those displayed for the 5mC and 5hmC content as measured by either the ELISA (r = 0.72, p =<0.001), or the GLUMA (r = -0.79, p =<0.001) method.

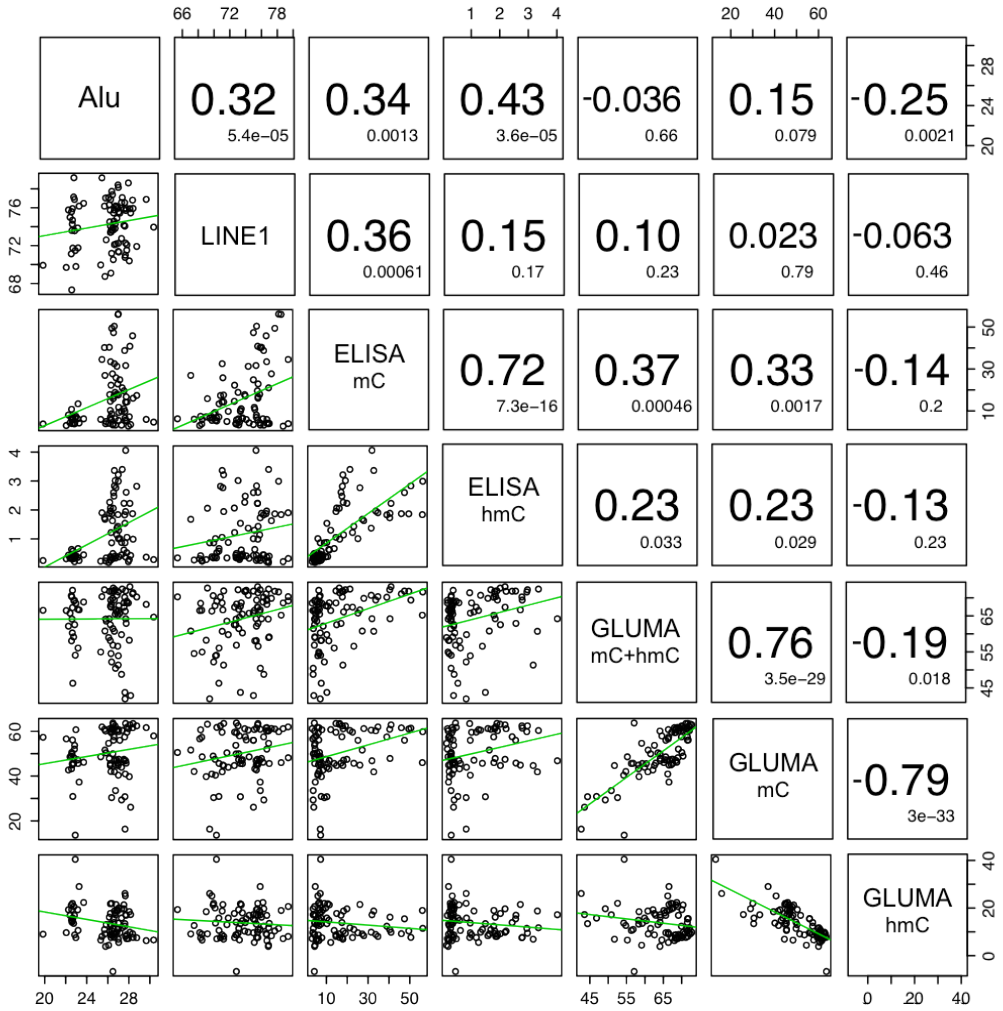


Figure 3.5 Scatterplot & correlation matrix of methylation by the different methods [lower panel: scatterplots; green line: line of least squares; upper panel: Pearson correlation coefficient, with p value in smaller text]

3.6.2 Global methylation by age

Unadjusted linear regression analysis of data from all brain areas by each method in turn revealed that age was significantly associated with both LINE1 and GLUMA 5hmC methylation levels. No significant associations of age with any of the other measures were observed. LINE1 methylation levels significantly decrease with increasing age (β = -0.05 % methylation/year, SE = 0.02, p = 0.042) (Figure 3.6), whereas GLUMA 5hmC levels significantly increase with age (β = 0.14 % methylation/year, SE = 0.04, p = 0.001) (Figure 3.7).

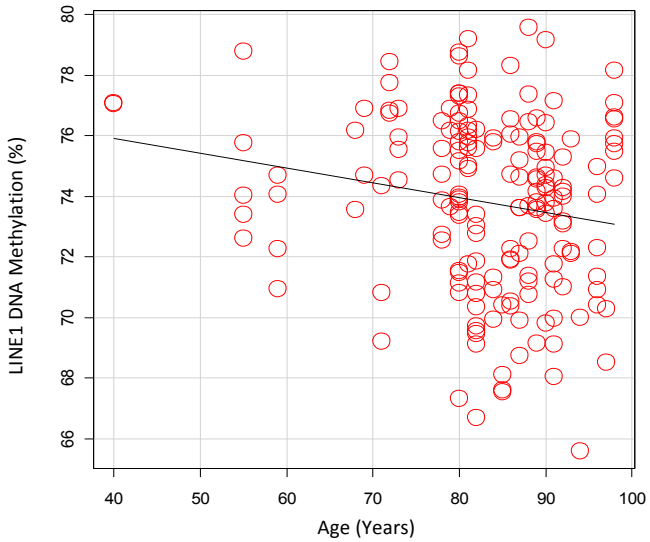


Figure 3.6 Scatterplot of LINE1 methylation levels by age. [black line: line of least squares]

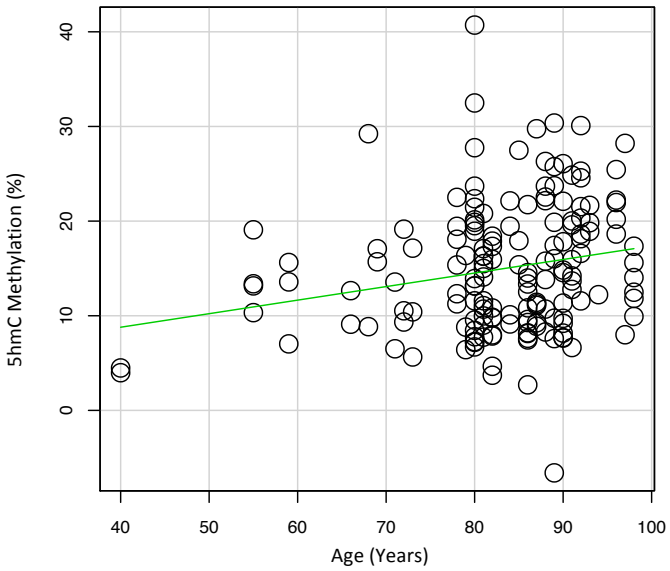
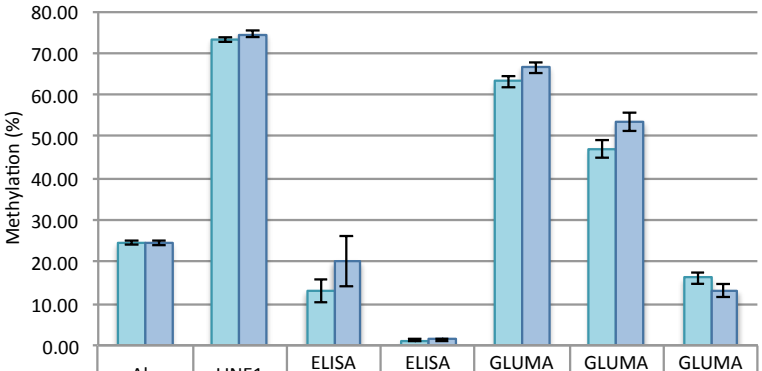


Figure 3.7 Scatterplot of 5hmC methylation levels by age. [green line: line of least squares]

3.6.3 Global methylation by sex

For each measure of DNA methylation, the mean levels of methylation across all brain areas, show a trend for higher levels of methylation in males compared to females with the exception of Alu repeat methylation, and GLUMA 5hmC which exhibits a trend in the opposite direction (Figure 3.8). To test if these sex differences were significant unadjusted linear regressions were run for each measure of DNA methylation in turn.



	Alu	LINE1	ELISA mC	ELISA hmC	GLUMA mC+hmC	GLUMA mC	GLUMA hmC
Female	24.59	73.18	13.02	1.05	63.10	46.98	16.11
Male	24.56	74.48	20.08	1.31	66.58	53.45	13.13

Female n=	37	37	23	23	38	38	38
Male n=	18	18	13	13	22	21	21

Figure 3.8 Overall average methylation levels by assay and sex
[whiskers: 95% confidence interval; n per group tabulated at bottom of figure]

Though Alu and LINE1 repeat methylation was not significantly associated with sex, when looking at all males and females regardless of disease status (β = -0.08, SE = 0.20, p = 0.67 and β = -0.60, SE = 0.6, p = 0.35 respectively), there are significant differences within the normal control samples (Figures 3.9 and 3.10). Alu DNA methylation differed by 0.64% between control males and females (β = -0.64, SE = 0.21, p = 0.007), whereas LINE1 methylation differed significantly by 2.29% (β = -2.29, SE = 0.97, p = 0.03), with males having more methylation in both instances.

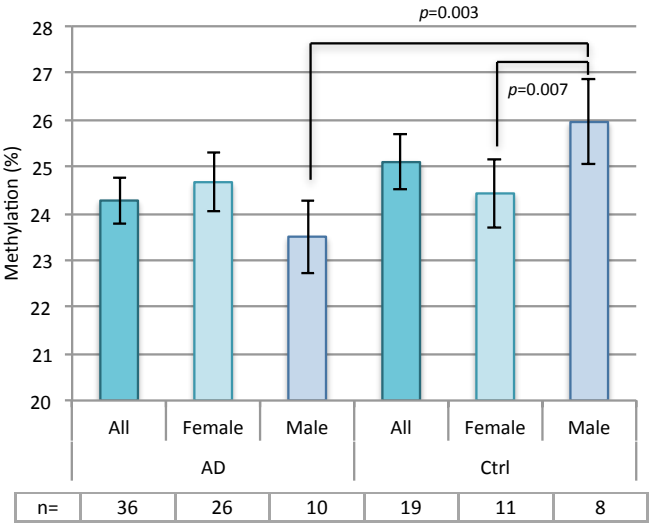


Figure 3.9 Overall average Alu methylation by disease status and sex
[AD: Alzheimer's disease; Ctrl: normal controls; whiskers: 95% confidence interval; n per group tabulated at bottom of figure]

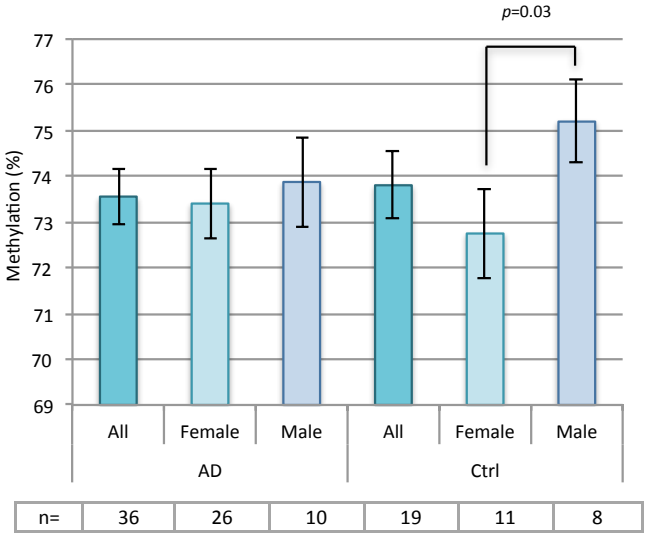
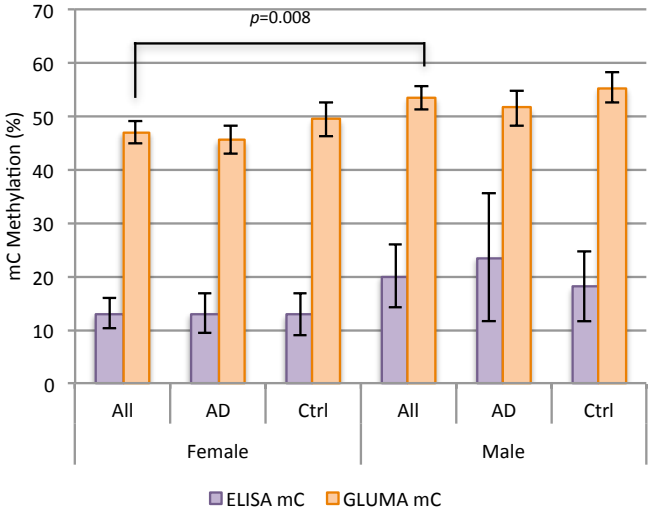


Figure 3.10 Overall average LINE1 methylation by disease status and sex
[AD: Alzheimer's disease; Ctrl: normal control; whiskers: 95% confidence interval; n per group tabulated at bottom of figure]

Total GLUMA 5mC+5hmC was significantly associated with sex (β = -2.77, SE = 1.3, *p* = 0.039), with males having on average 2.77% more total content methylation compared to females. More specifically, the content of GLUMA 5mC differed significantly with males having 5.23% more 5mC methylation compared to females (β = -5.23, SE = 1.9, *p* = 0.008) (Figure 3.11). While not statistically significant a similar trend for these sex differences was also seen for the ELISA 5mC measure (β = -6.68, SE = 4.96, p = 0.19) as shown alongside the GLUMA 5mC data in Figure 3.11. While no other age or sex differences were observed all the following analyses were controlled for age and sex, along with ApoE genotype and batch effects where necessary.

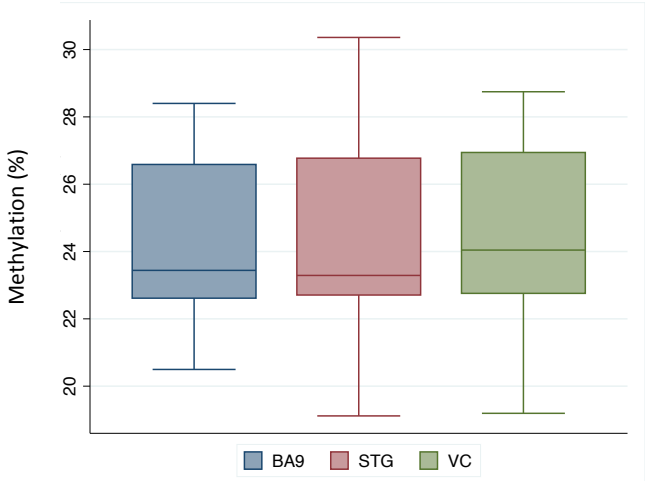


ELISA mC n=	23	15	8	13	5	8
GLUMA mC n=	38	24	14	21	9	12

Figure 3.11 Overall average measured 5mC methylation by sex and disease status
[All: cases and controls; AD: Alzheimer's disease; Ctrl: normal control; whiskers: 95% confidence interval; n per group tabulated at bottom of figure]

3.6.4 Alu & LINE1 consensus CpG global methylation

The average level of Alu CpG DNA methylation (5mC+5hmC) detected in the brain was 25.9%; with average LINE1 CpG methylation in the brain much higher at 74.2%. The average levels of each measure across the three brain areas tested are also very similar (Figures 3.12-3.13)



n=	40	56	55

Figure 3.12 Average methylation levels for Alu CpGs by tissue
[BA9: Brodmann Area 9; STG: Superior Temporal Gyrus; VC: Visual Cortex; whiskers: 95% confidence interval; n per group tabulated at bottom of figure]

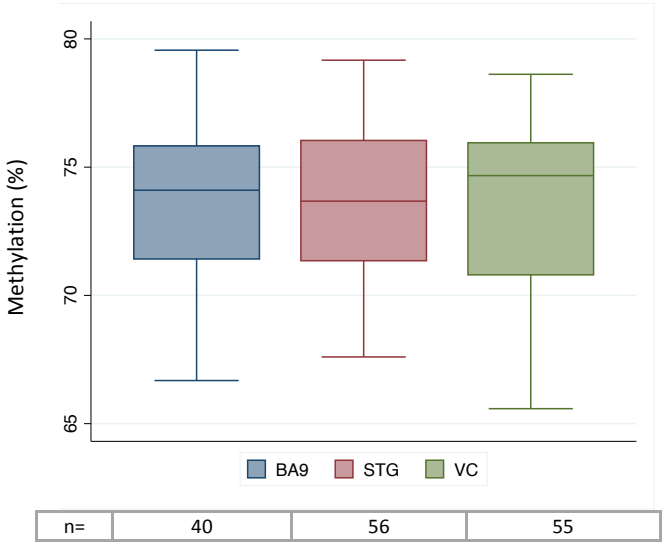


Figure 3.13 Average methylation levels for LINE1 CpGs by tissue
[BA9: Brodmann Area 9; STG: Superior Temporal Gyrus; VC: Visual Cortex; whiskers: 95% confidence interval; n per group tabulated at bottom of figure]

Methylation values for Alu elements show a similar profile across the three different brain regions and by AD status (Figure 3.14). Comparing the average case verses control methylation levels for Alu at the each brain region shows a trend for a decline in AD brain compared to controls. The difference is more pronounced at the STG and the VC areas, whereas the BA9 area shows less of a difference if any (Figure 3.14). Association of either, AD status or Braak staging with Alu methylation was sought by both linear regressions for each brain area in turn and mixed models for all the data from the three brain areas, using the VC as baseline. However no significant associations were identified. Results from the Alu methylation mixed models using AD status and Braak staging are presented in Tables 3.4-3.5.

Linear regression of the overall average Alu methylation levels in brain shows a significant decline of 0.8% in AD brain compared to controls, for male samples only (β = -0.80, SE = 0.25, p = 0.003) (Figure 3.9). On testing individual brain areas (by linear regression) for an association of AD status (in males) with Alu methylation levels there were significantly associated with lower levels compared to controls at the VC and STG areas (β = -1.31, SE = 0.59, p = 0.038 and β = -1.05, SE = 0.48, p = 0.04 respectively), but not the BA9 area (β = -0.05, SE = 0.42, p = 0.9).

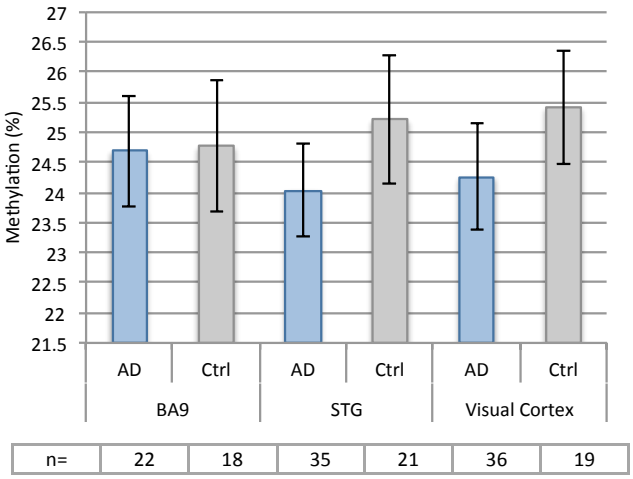


Figure 3.14 Average methylation levels for Alu CpGs by tissue & disease status
[BA9: Brodmann Area 9; STG: Superior Temporal Gyrus; VC: Visual Cortex; AD: Alzheimer's disease; Ctrl: normal control; whiskers: 95% confidence interval; n per group tabulated at bottom of figure]

Variable		Coefficient	Std. Err.	Z	*p*	95% CI	
AD status		-0.239	0.204	-1.170	0.240	-0.638	0.160
Sex		-0.265	0.184	-1.440	0.150	-0.625	0.095
Age at death		0.008	0.010	0.800	0.425	-0.012	0.027
ApoE genotype		0.004	0.017	0.220	0.823	-0.029	0.037
Brain area:	STG	-0.121	0.177	-0.680	0.496	-0.468	0.226
	BA9	-0.033	0.231	-0.140	0.888	-0.485	0.420

Table 3.4 Mixed model output of variables on Alu methylation using AD status as disease marker
[STG: Superior Temporal Gyrus; *: significant variable]

Variable		Coefficient	Std. Err.	Z	*p*	95% CI	
Braak staging		-0.014	0.045	-0.310	0.759	-0.103	0.075
Sex		-0.316	0.193	-1.640	0.101	-0.694	0.062
Age at death		0.006	0.010	0.590	0.555	-0.014	0.026
ApoE genotype		-0.001	0.017	-0.030	0.974	-0.034	0.033
Brain area:	STG	-0.154	0.180	-0.850	0.393	-0.508	0.200
	BA9	-0.091	0.235	-0.390	0.697	-0.551	0.368

Table 3.5 Mixed model output of variables on Alu methylation using Braak staging as disease marker
[STG: Superior Temporal Gyrus; *: significant variable]

Similar differences in methylation at LINE1 sequences can be seen at the visual cortex, however the STG brain area shows no difference and the BA9 area shows a slight increase in methylation for AD samples compared to controls (Figure 3.15). The trend for a decline in DNA methylation of AD samples compared to controls is more apparent in Alu CpG methylation across the three brain areas, than it is at LINE1 CpG methylation. Association of either AD status or Braak staging with LINE1 methylation was sought by both linear regressions for each brain area in turn and

mixed models for all the data from the three brain areas, using the VC as baseline. However no significant associations were identified. Results from the LINE1 methylation mixed models using AD status and Braak staging are presented in Tables 3.6-3.7.

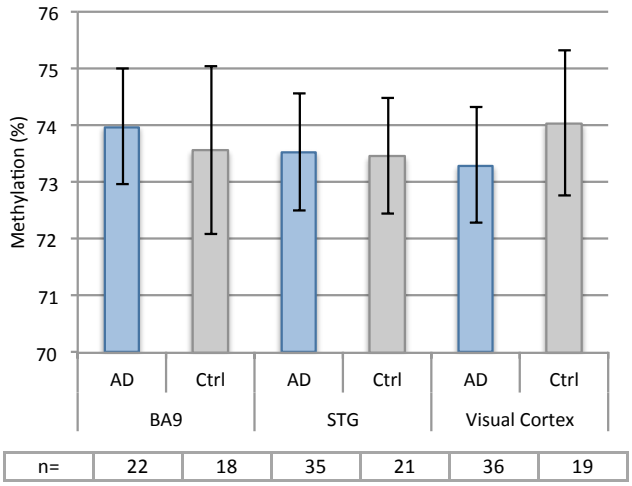


Figure 3.15 Average methylation levels for LINE1 CpGs by tissue & disease status
[BA9: Brodmann Area 9; STG: Superior Temporal Gyrus; VC: Visual Cortex; AD: Alzheimer's disease; Ctrl: normal control; whiskers: 95% confidence interval; n per group tabulated at bottom of figure]

Variable		Coefficient	Std. Err.	Z	*p*	95% CI	
AD status		0.206	0.804	0.260	0.798	-1.370	1.782
Sex		0.130	0.659	0.200	0.843	-1.162	1.422
Age at death		-0.061	0.035	-1.750	0.081	-0.130	0.008
ApoE genotype		-0.054	0.064	-0.860	0.392	-0.179	0.070
Brain area:	STG	-0.127	0.361	-0.350	0.725	-0.834	0.581
	BA9	-0.119	0.455	-0.260	0.794	-1.010	0.772

Table 3.6 Mixed model output of variables on LINE1 methylation using AD status as disease marker
[STG: Superior Temporal Gyrus; *: significant variable]

Variable		Coefficient	Std. Err.	Z	*p*	95% CI	
Braak staging		0.031	0.178	0.180	0.860	-0.318	0.380
Sex		-0.220	0.704	-0.310	0.755	-1.599	1.160
Age at death		-0.056	0.036	-1.550	0.122	-0.128	0.015
ApoE genotype		-0.058	0.065	-0.890	0.375	-0.186	0.070
Brain area:	STG	-0.193	0.368	-0.520	0.600	-0.915	0.529
	BA9	-0.205	0.463	-0.440	0.658	-1.111	0.702

Table 3.7 Mixed model output of variables on LINE1 methylation using Braak staging as disease marker
[STG: Superior Temporal Gyrus; *: significant variable]

3.6.5 5mC/5hmC global methylation by ELISA

Recently available ELISA kits from Epigentek, were used to quantify global 5mC and 5hmC. Initial testing showed highly variable sample preparation specific variation, as seen in Figure 3.16, seven samples were tested in duplicate across four different tissues; not only was the 5mC measured across the different tissues questionably variable, there was also notable variation between each individual sample duplicate. Checking the genomic DNA samples for RNA contamination and treatment with RNase reduced this variability and thus indicated there was possibly some cross reactivity with RNA 5-methylcytosine species. Subsequently all samples tested and presented in the remainder of this section were RNase treated, prior to testing.

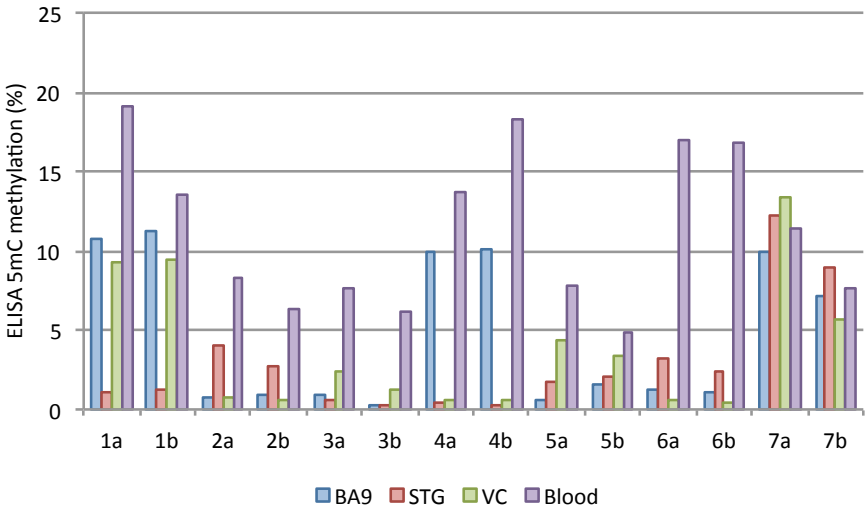


Figure 3.16 Initial ELISA mC results across four tissues
[BA9: Brodmann Area 9; STG: Superior Temporal Gyrus; VC: Visual Cortex; a and b: duplicate samples]

The average level of 5mC methylation detected in the brain was 15.9% whilst 5hmC methylation was at a much lower level of 1.2% (total 5mC+5hmC = 17.1%). Results from both the 5mC and 5hmC ELISA assays show very similar profiles, with the BA9 area having slightly higher levels, compared to the STG and VC areas (Figures 3.17-3.18).

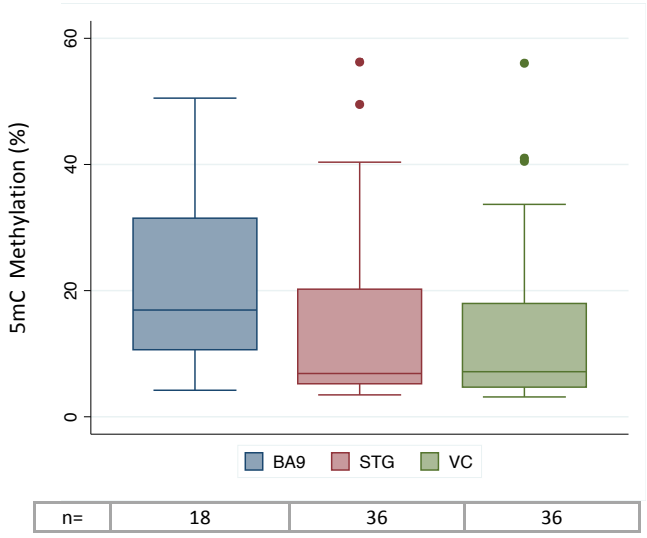


Figure 3.17 Average methylation levels for ELISA 5mC by tissue
[BA9: Brodmann Area 9; STG: Superior Temporal Gyrus; VC: Visual Cortex; whiskers: 95% confidence interval; n per group tabulated at bottom of figure]

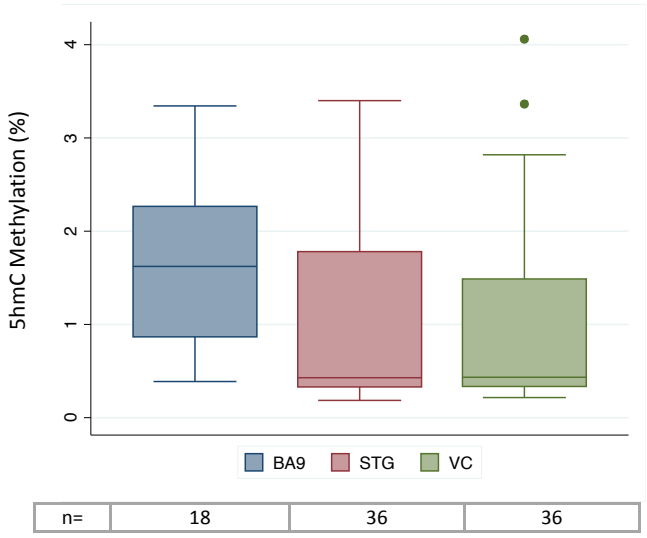


Figure 3.18 Average methylation levels for ELISA 5hmC by tissue
[BA9: Brodmann Area 9; STG: Superior Temporal Gyrus; VC: Visual Cortex; whiskers: 95% confidence interval; n per group tabulated at bottom of figure]

For both the 5mC and 5hmC measures, looking at the AD pathology affected brain region of the STG, AD samples show a trend for a decline in average global methylation compared to controls, whereas the inverse trend is seen at the BA9 brain area (Figures 3.19-3.20). Average levels of both the 5mC and 5hmC are much the same in the visual cortex an AD pathology unaffected brain region.

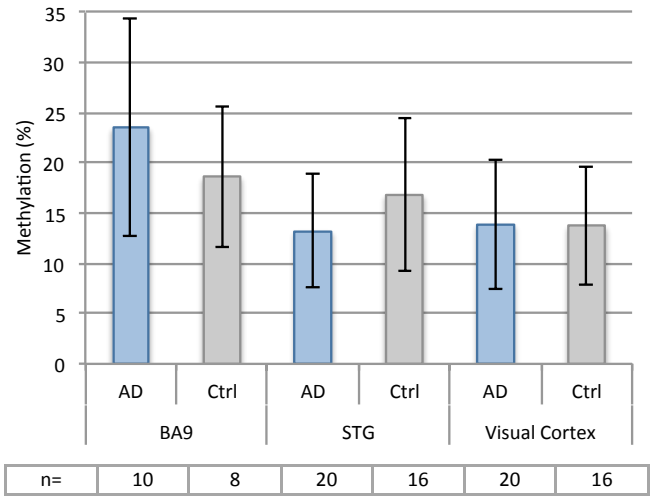


Figure 3.19 Average methylation levels for ELISA 5mC by tissue & disease status
[BA9: Brodmann Area 9; STG: Superior Temporal Gyrus; VC: Visual Cortex; AD: Alzheimer's disease; Ctrl: normal control; whiskers: 95% confidence interval; n per group tabulated at bottom of figure]

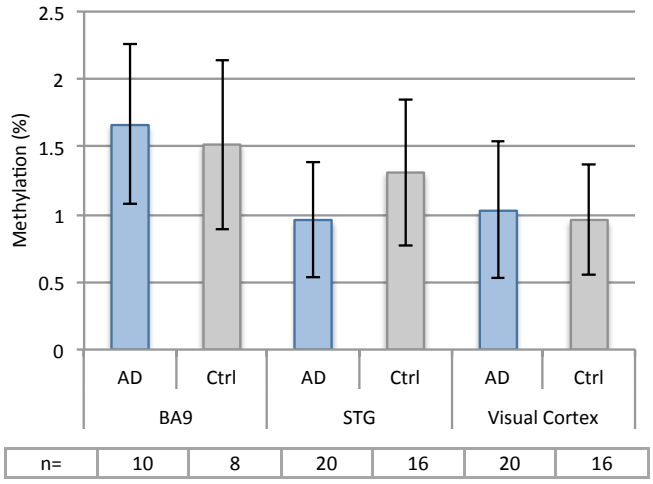


Figure 3.20 Average methylation levels for ELISA 5hmC by tissue & disease status
[BA9: Brodmann Area 9; STG: Superior Temporal Gyrus; VC: Visual Cortex; AD: Alzheimer's disease; Ctrl: normal control; whiskers: 95% confidence interval; n per group tabulated at bottom of figure]

Association of either, AD status or Braak staging with each of the ELISA methylation measures was formally tested by both linear regressions for each brain area in turn and mixed models for all the data from the three brain areas, using the VC as baseline. However no significant associations were identified. Results from the ELISA 5mC and 5hmC mixed models using AD status and Braak staging are summarised in Tables 3.8-3.11.

Variable		Coefficient	Std. Err.	Z	*p*	95% CI	
AD status		-0.864	5.156	-0.170	0.867	-10.969	9.241
Sex		-3.087	3.526	-0.880	0.381	-9.997	3.823
Age at death		0.157	0.206	0.760	0.445	-0.247	0.562
ApoE genotype		-0.204	0.508	-0.400	0.688	-1.198	0.791
Brain area:	STG	1.021	1.518	0.670	0.501	-1.954	3.996
	BA9	0.977	2.172	0.450	0.653	-3.280	5.235

Table 3.8 Mixed model output of variables on ELISA 5mC methylation using AD status as disease marker [STG: Superior Temporal Gyrus; *: significant variable]

Variable		Coefficient	Std. Err.	Z	*p*	95% CI	
Braak staging		-0.262	1.128	-0.230	0.816	-2.474	1.950
Sex		-3.801	3.965	-0.960	0.338	-11.572	3.970
Age at death		0.181	0.216	0.840	0.401	-0.242	0.604
ApoE genotype		-0.181	0.505	-0.360	0.719	-1.170	0.808
Brain area:	STG	1.209	1.609	0.750	0.452	-1.944	4.362
	BA9	1.100	2.226	0.490	0.621	-3.264	5.463

Table 3.9 Mixed model output of variables on ELISA 5mC methylation using Braak staging as disease marker [STG: Superior Temporal Gyrus; *: significant variable]

Variable		Coefficient	Std. Err.	Z	*p*	95% CI	
AD status		-0.265	0.367	-0.720	0.470	-0.984	0.454
Sex		-0.171	0.275	-0.620	0.533	-0.709	0.367
Age at death		0.013	0.015	0.840	0.399	-0.017	0.042
ApoE genotype		0.032	0.036	0.880	0.381	-0.039	0.103
Brain area:	STG	0.107	0.134	0.800	0.422	-0.155	0.369
	BA9	0.118	0.172	0.690	0.493	-0.219	0.454

Table 3.10 Mixed model output of variables on ELISA 5hmC methylation using AD status as disease marker [STG: Superior Temporal Gyrus; *: significant variable]

Variable		Coefficient	Std. Err.	Z	*p*	95% CI	
Braak staging		-0.117	0.078	-1.510	0.131	-0.269	0.035
Sex		-0.237	0.295	-0.800	0.422	-0.815	0.341
Age at death		0.019	0.015	1.270	0.205	-0.010	0.049
ApoE genotype		0.040	0.035	1.140	0.256	-0.029	0.108
Brain area:	STG	0.125	0.141	0.880	0.377	-0.152	0.401
	BA9	0.133	0.177	0.750	0.452	-0.213	0.479

Table 3.11 Mixed model output of variables on ELISA 5hmC methylation using Braak staging as disease marker [STG: Superior Temporal Gyrus; *: significant variable]

3.6.6 Glucosylation LUminometric Methylation Assay (GLUMA)

Given that the variance of the average methylation for the sample groups tested was considerable for both the repeat elements Alu/LINE1 and ELISA, a modified version of the LUMA assay was developed that would allow for the simultaneous measurement of both 5mC and 5hmC methylation levels in a single assay.

During the initial testing/development a small number of blood samples and three of mouse ES cell lines were measured to ascertain the validity of this assay, as they were expected to exhibit distinct levels of 5hmC compared to brain tissue. As expected the blood samples showed the lowest levels of 5hmC measured across the different tissues and the mouse ES cells exhibited levels that in intermediate to those observed for blood and brain tissues (Figure 3.21).

Average DNA methylation in all brain samples tested by the GLUMA method was 64.1% for total methylation (5mC+5hmC), 50.1% for 5mC and 13.9% for 5hmC. Average levels of both total methylation and 5mC show a trend for a decline in levels from the BA9 area (66.6% and 52.3%) the STG (65.7% and 51.4%) to and VC (61.8% and 45.3%). Though a similar trend persists in the opposite direction for the 5hmC levels, with the VC having 16.5%, the STG and BA9 areas have the same mean level at 14.3% (Figure 3.21).

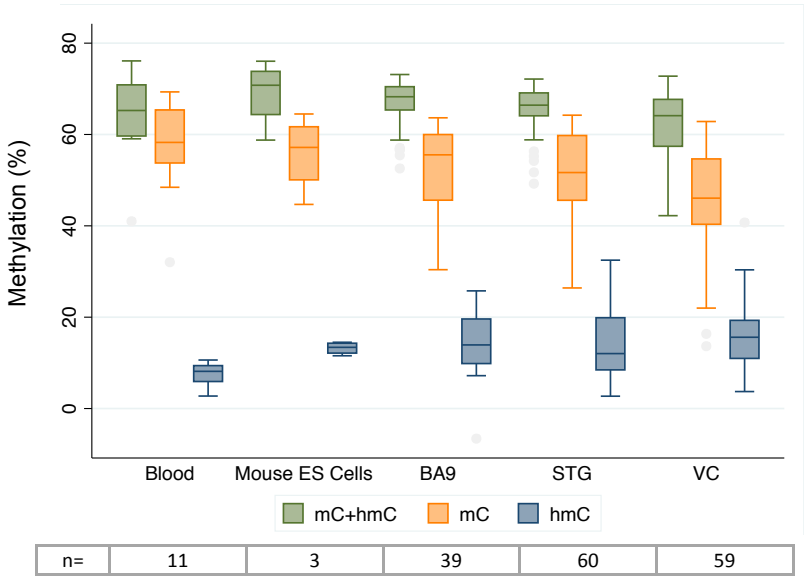


Figure 3.21 Average methylation levels for GLUMA results across four tissues
[BA9: Brodmann Area 9; STG: Superior Temporal Gyrus; VC: Visual Cortex; whiskers: 95% confidence interval; n per group tabulated at bottom of figure]

As mentioned earlier the correlation of measured 5hmC and 5mC levels by the GLUMA method show an inverse relationship (r = -0.79, p =<0.001). This would be consistent with the notion

that 5hmC serves as an intermediate form for de-methylation of 5mC to standard cytosine (Figure 3.22).

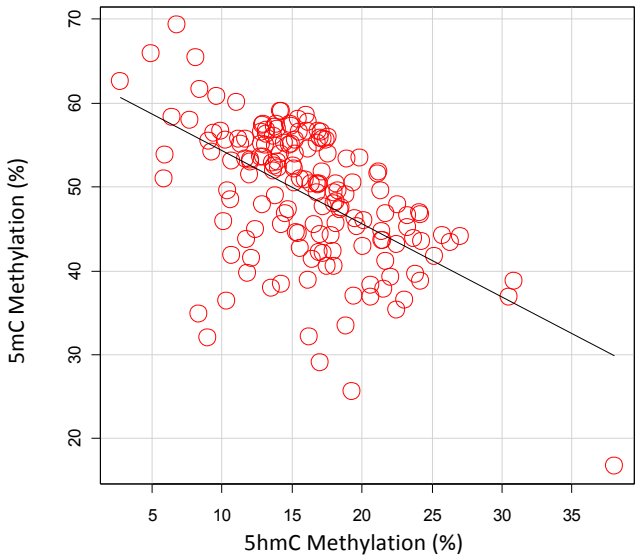
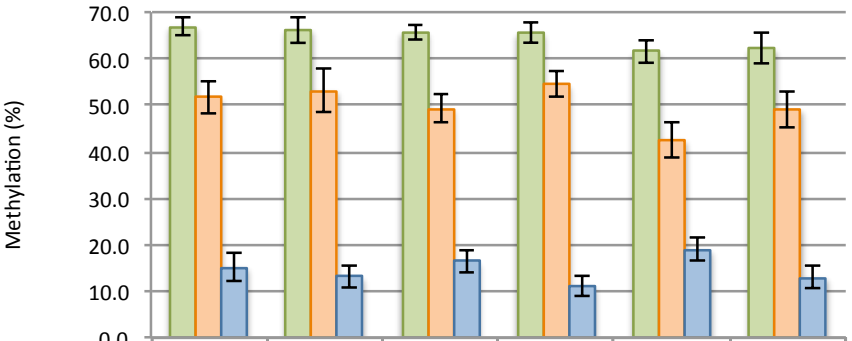


Figure 3.22 Scatterplot of 5hmC methylation against 5mC methylation [black line: line of least squares]

While the average levels of total methylation content are similar for cases and controls in the three brain areas, AD samples show reduced levels of 5mC compared to control samples, with the inverse relationship for 5hmC levels (Figure 3.23). These differences are more apparent at the STG and VC areas compared to the BA9 area. On average both the 5mC and 5hmC levels for AD and controls at the STG differed by ~5%. With a 5mC difference of 6.5% and a 5hmC difference of 5.9% between AD and controls at the visual cortex (Figure 3.23).



	AD	Ctrl	AD	Ctrl	AD	Ctrl
	BA9		STG		VC	
mC+hmC%	66.9	66.2	65.7	65.5	61.5	62.1
mC%	51.7	53.1	49.4	54.4	42.6	49.1
hmC%	15.2	13.1	16.3	11.1	18.9	13.0

n=	22	17	34	26	33	26

Figure 3.23 Average GLUMA measured methylation by disease status and brain area [AD: Alzheimer's disease; Ctrl: normal control; whiskers: 95% confidence interval; n per group tabulated at bottom of figure]

Comparing the levels of 5mC measured by the ELISA and GLUMA methods (Figures 3.19 and 3.23), the same trend is seen at STG brain area according to AD status, where AD samples present with lower levels of 5mC compared to controls. However, while the GLUMA continues to shows the same trend across the three tissues, the ELISA shows the opposite trend at the BA9 area and no difference in levels at the VC. For the 5hmC measured by the two methods, the same trends are seen for the ELISA 5hmC as the ELISA 5mC, while the GLUMA 5hmC levels are higher for AD samples compared to controls at each brain area (Figures 3.20 and 3.23).

On testing individual brain areas (by linear regression) for an association of either, AD status or Braak staging with each of the three methylation measures generated by the GLUMA assay, both AD status and a higher Braak stage were significantly associated with higher 5hmC levels compared to controls at the VC area ($\beta = 4.13$, SE = 1.96, $p = 0.04$ and $\beta = 0.99$, SE = 0.45, $p = 0.03$ respectively), whereas only Braak stage was significantly associated with 5hmC levels at the STG brain area ($\beta = 1.10$, SE = 0.49, $p = 0.026$). Although the 5mC measure also shows a trend for association with Braak staging at the STG area ($\beta = -1.33$, SE = 0.67, $p = 0.053$), this was not statistically significant, as were the results for all the other remaining combinations of methylation measure and brain area.

The same mixed model approach as used above for the other measures of global methylation was applied to each GLUMA methylation measure in turn, to test for any significant differences between brain areas or associations with AD status or Braak staging. Consistently, both the levels of methylation at the STG and BA9 areas are significantly different from the baseline VC, for each measure in all the models tested (Tables 3.12-3.17). No significant associations of either of the disease indicators with the total 5mC+5hmC methylation content were observed, however as noted, total methylation content at the STG and the BA9 differed significantly from the VC by 3.7% and 5% respectively (Tables 3.12 and 3.13).

Variable		Coefficient	Std. Err.	Z	*p*	95% CI	
AD status		0.944	1.340	0.700	0.481	-1.682	3.571
Sex		-2.041	1.146	-1.780	0.075	-4.287	0.205
Age at death		0.054	0.063	0.850	0.396	-0.070	0.178
ApoE genotype		-0.048	0.112	-0.430	0.668	-0.267	0.171
Brain area:	*STG	3.677	0.606	6.060	<0.001	2.488	4.865
	*BA9	5.091	0.769	6.620	<0.001	3.583	6.599

Table 3.12 Mixed model output of variables on GLUMA 5mC+5hmC methylation using AD status as disease marker [STG: Superior Temporal Gyrus; *: significant variable]

Variable		Coefficient	Std. Err.	Z	*p*	95% CI	
Braak staging		-0.366	0.311	-1.180	0.239	-0.976	0.244
Sex		-2.241	1.227	-1.830	0.068	-4.646	0.165
Age at death		0.103	0.067	1.520	0.127	-0.029	0.235
ApoE genotype		0.011	0.115	0.090	0.927	-0.215	0.236
Brain area:	*STG	3.750	0.630	5.960	<0.001	2.517	4.984
	*BA9	5.120	0.807	6.340	<0.001	3.537	6.702

Table 3.13 Mixed model output of variables on GLUMA 5mC+5hmC methylation using Braak staging as disease marker [STG: Superior Temporal Gyrus; *: significant variable]

For the measure of 5mC, while AD status is not significantly associated with 5mC levels (β = -1.487, SE = 2.119, *p* = 0.483) there is a significant association of Braak staging with 5mC levels (β = -1.09, SE = 0.478, *p* = 0.027), on average with each increasing level of Braak stage there is a decrease of 1.1% 5mC. Also the STG and BA9 show significantly higher levels of 5mC methylation (5.7% and ~8% respectively) compared to the VC. Results from the GLUMA 5mC methylation mixed models using AD status and Braak staging are presented in Tables 3.14-3.15.

Variable		Coefficient	Std. Err.	Z	*p*	95% CI	
AD status		-1.487	2.119	-0.700	0.483	-5.640	2.665
*Sex		-3.971	1.868	-2.130	0.034	-7.633	-0.309
Age at death		-0.067	0.097	-0.690	0.491	-0.258	0.124
ApoE genotype		-0.091	0.175	-0.520	0.604	-0.433	0.251
Brain area:	*STG	5.662	1.196	4.740	<0.001	3.319	8.005
	*BA9	7.527	1.422	5.290	<0.001	4.739	10.314

Table 3.14 Mixed model output of variables on GLUMA 5mC methylation using AD status as disease marker [STG: Superior Temporal Gyrus; *: significant variable]

Variable		Coefficient	Std. Err.	Z	*p*	95% CI	
*Braak staging		-1.059	0.478	-2.210	0.027	-1.995	-0.122
*Sex		-4.934	1.935	-2.550	0.011	-8.726	-1.143
Age at death		0.014	0.101	0.140	0.887	-0.183	0.211
ApoE genotype		-0.038	0.174	-0.220	0.826	-0.380	0.303
Brain area:	*STG	5.785	1.229	4.710	<0.001	3.376	8.194
	*BA9	7.918	1.470	5.390	<0.001	5.036	10.799

Table 3.15 Mixed model output of variables on GLUMA 5mC methylation using Braak staging as disease marker [STG: Superior Temporal Gyrus; *: significant variable]

Both AD status and Braak staging are significantly associated with 5hmC levels. AD patients on average have 2.68% more 5hmC compared to controls (β = 2.68, SE = 1.35, *p* = 0.048), whereas on average with each increasing level of Braak stage there is an increase of 0.85% 5hmC (β = 0.834, SE = 0.299, *p* = 0.005). Also the STG and BA9 show significantly lower levels of 5hmC methylation (1.9% and ~2.5% respectively) compared to the VC. Results from the GLUMA 5hmC methylation mixed models using AD status and Braak staging are presented in Tables 3.16-3.17.

Variable		Coefficient	Std. Err.	Z	*p*	95% CI	
*AD status		2.682	1.354	1.980	0.048	0.029	5.335
Sex		1.787	1.209	1.480	0.139	-0.583	4.157
Age at death		0.102	0.061	1.660	0.097	-0.019	0.222
ApoE genotype		0.028	0.111	0.250	0.800	-0.190	0.246
Brain area:	*STG	-1.910	0.904	-2.110	0.035	-3.683	-0.137
	*BA9	-2.432	1.055	-2.310	0.021	-4.499	-0.365

Table 3.16 Mixed model output of variables on GLUMA 5hmC methylation using AD status as disease marker [STG: Superior Temporal Gyrus; *: significant variable]

Variable		Coefficient	Std. Err.	Z	*p*	95% CI	
*Braak staging		0.834	0.299	2.790	0.005	0.248	1.420
*Sex		2.587	1.219	2.120	0.034	0.197	4.977
Age at death		0.062	0.062	1.000	0.315	-0.059	0.184
ApoE genotype		0.014	0.109	0.130	0.897	-0.199	0.227
Brain area:	*STG	-1.853	0.909	-2.040	0.042	-3.635	-0.070
	*BA9	-2.733	1.061	-2.580	0.010	-4.814	-0.653

Table 3.17 Mixed model output of variables on GLUMA 5hmC methylation using Braak staging as disease marker [STG: Superior Temporal Gyrus; *: significant variable]

Within these mixed models where disease marker was significantly associated with methylation, a interaction term of disease marker and brain area was tested, to asses whether the relationship between disease marker and methylation differs according to brain area, however such a interaction was not significant.

3.7 Discussion

While there are reports of both hypo- and hyper-methylation with aging in the literature (Rath & Kanungo 1989; Kim et al. 2009; Pogribny & Vanyushin 2010), we observed a significant association of a 0.05% decrease in LINE1 methylation per year. In line with previous findings in mouse brain (Münzel et al. 2010; Song et al. 2010; Chen et al. 2012) an association of increasing GLUMA 5hmC methylation with increasing age was also observed (0.14% per year). Though these results are based only on data from individuals at death and not from the same individual over different time points, they give an indication as to changes in global methylation over time.

In line with previous studies (Bollati et al. 2011; El-Maarri et al. 2011) overall males compared to females had significantly elevated levels of DNA methylation. Clear sex specific differences were observed within the normal controls at Alu and LINE1 consensus CpG sites. The same differences extended to all the samples for both the ELISA and GLUMA measures of 5mC, though this difference at the ELISA 5mC was not significant. This observation of preserved differences in Alu and LINE1 methylation within the control samples compared to cases could

possibly be viewed as a distortion of the normal methylation pattern across repeat sequences in AD brain, or just as likely maybe due the presence of extensive neurodegeneration in AD which may lead to the measurement of altered global methylation compared to non-AD brains (see below).

Although there were trends for differences in the average levels of DNA methylation of AD samples compared to controls as measured by Alu, ELISA 5mC and ELISA 5hmC none of these are significant. However Alu methylation was significantly decreased in male AD brain by 0.8% compared to male controls and though this is similar to that observed by Bollati et al. (2011) for both sexes in AD peripheral lymphocytes, their analysis did not reach statistical significance once adjusted for age and sex. While the observation of Alu methylation difference in male brain samples reported here is based on a small numer of samples (AD=10, Ctrl=8) with a measure across three brain areas per sample the analysis was adjusted for age, sex and ApoE genotype. Even though the measured change in Alu methylation is very small, given that Alu sequences account for around 26% of CpGs in the genome (Rollins et al. 2006) and are enriched in active domains of the genome (The ENCODE Project, 2007), a small change could possibly affect a large number of potentially important CpG sites throughout the genome.

Furthermore Bollati et al. (2011) observed a significant increase in LINE1 methylation in AD peripheral lymphocytes compared to controls, whereas the current data on post-mortem brain data show no significant differences in LINE1 methylation between cases and controls. Though these differences in results between studies may reflect real tissue specific changes, both the current post-mortem brain data presented here and those for peripheral lymphocytes are based on small sample sizes, vunerable to sample bias and thus the possibliity of chance findings can not be ruled out (Bollati et al. 2011). Likewise the cell composition of the starting material/tissue used in the analysis could also influence the results. For instance, Wu et al. (2011) have demonstrated that both assay type and cell composition of white-blood cells can affect the results of global methylation assays.

Simultaneous measurement of both 5mC and 5hmC methylation was successfully achieved with a modified version of the LUMA assay (GLUMA). Whereby significant differences were measured in both 5mC and 5hmC between the AD pathology-affected and unaffected brain regions and either AD status or Braak staging were significantly associated with 5hmC methylation as measured by the GLUMA assay. Also as Braak staging was noted to be significantly associated with 5mC levels whereas AD status was not, therefore Braak staging may

provide a more sensitive scale on which disease and brain pathology is defined, as opposed to the dichotomous grouping of AD status.

However the differences between brain regions was not in the direction that we expected with lower levels of 5mC and higher levels of 5hmC in the visual cortex compared to the STG region. This may possibly be due to that fact that these measurements are made on DNA extracted from brain tissue homogenates with unknown cell composition and the fact that the diseased brain has a loss of neurons relative to normal brain.

One option to over come this would be to use cell populations selectively collected using a method such as Fluorescence-Activated Cell Sorting (FACS) or Laser Capture Microdissection (LCM). While the cells collected using these methods provide only a small number of cells and thus low yields from DNA extraction which may not be suitable to the requirements of the methods used in this thesis, optimized protocols of a genome wide -CCGG- methylation sequencing method, Reduced Representation Bisulfite Sequencing (RRBS) have been reported to have used as little as 30-100 ng of DNA (Meissner et al. 2008; Boyle et al. 2012; Gu et al. 2010), which maybe more applicable to the quantities of DNA obtained via FACS or LCM.

Nonetheless the levels of 5mC are lower and 5hmC levels higher in AD brain compared to controls, across both AD pathology affected (STG) and unaffected brain (VC) areas, highlighting the fact that though the changes in levels observed in the STG area maybe the result of changes in cell composition, similarly striking differences are also observed in the VC brain area that is expected to be spared of such changes. Although the VC was used an AD-pathology spared brain area, this is probably an over simplification as such areas may show NFTs and amyloid-β plaques in more advanced disease. Cerebellum, however, is relatively resistant to neuronal damage with little or no tangle formation, and may be a better choice for a brain area that is spared of AD pathology. The methylation changes observed at the STG and VC brain area are suggestive of advanced disease and may reflect an affect of the oxidative and the inflammatory processes active in the degenerating AD brain. A major limitation of this study and others looking at molecular changes in AD is the lack of longitudinal samples from early stage AD to advanced disease and though post-mortem brain tissues samples, help to understand disease mechanisms, they only provide a snap-shot of the end of the disease process, which may not necessarily reflect the mechanisms responsible for disease development (Bollati et al. 2011).

Interestingly, if only the total methylation (5mC+5hmC) at the -CCGG- sequence context throughout the genome was to be measured none of these significant differences at the

individual 5mC or 5hmC levels would be detected exemplifying the need to separate the different content of 5mC and 5hmC, particularly given their different proposed roles. The observation of no detectable difference in total methylation (5mC+5hmC) between cases and controls is in line with a previous study (Schwob et al. 1990) detecting no difference in methylation status of AD cases against controls at the -CCGG- sequence context, as this study did not differentiate 5mC from 5hmC.

Also we observed a inverse relationship in the levels of 5hmC and 5mC consistent with the notion that 5hmC serves as an intermediary form for de-methylation of 5mC. Given that 5hmC is generally enriched at active gene promoters; gene-bodies with increased transcription (Ficz et al. 2011; Khare et al. 2012) and lower levels at alternatively spliced transcripts (Khare et al. 2012), the observed differences in 5hmC levels between the brain of AD patients and controls may play a role in the neurodegenerative process. Furthermore gene ontology analysis of 5hmC genomic distribution in brain is suggestive of genes involved in neurodevelopment, synaptic plasticity and neurodegeneration (Khare et al. 2012; Song et al. 2011; Song et al. 2010).

While the highest level of 5hmC detected by the GLUMA assay is lower than those previously reported in human brain by the gold standard of LC-MS, ~0.34% compared to 0.7-1.2% of all cytosines (Kraus et al. 2012; Terragni et al. 2012), to the best of my knowledge this is the first time such changes in AD brain have been detected with this modification of the LUMA method. While these differences maybe attributable to the differing sensitivity of the methods used the differences measured here are likely more relative than absolute. In addition Kraus et al. (2012), comment on 5hmC being unevenly distributed in both neuronal and glial cells, suggesting their results unequivocally show that there is a wide variation in cellular 5hmC concentration. Thus this may even influence the differences in estimated levels along with the choice of brain area sampled in different studies.

At best the changes in DNA methylation levels observed here and in the literature are as small as 5% and are liable to the requirement of a large number of samples to reach any significance. The number of samples tested here using the ELISA method was small (N=35), given that there was uncertainty as to the accuracy of the kits used. Whereas the numbers of samples tested using the pyrosequencing and GLUMA methods were greater (N=56-60), here the limitation was the number of control brain samples available at the time of carrying out the experiments. Finally, though there was a concern that, due to the significantly smaller number of samples tested at the BA9 brain area, it would be better to exclude them from the mixed models tested, to allow for a more robust modeling of the STG and VC data alone. These analyses were also

performed excluding the BA9 area and there was virtually no difference in the results, and hence the BA9 data was included as presented here accounting for all the available data.

In summary the observed changes in global DNA methylation may reflect an affect of the oxidative and the inflammatory processes active in the degenerating Alzheimer's brain. Further investigations in a larger cohort of post-mortem brain samples with accompanying prospectively collected peripheral tissue are needed to clarify these preliminary observations and to better understand the role if any of Alu repeat element methylation and 5hmC in the neurodegenerative process.

SERT GENE PROMOTER METHYLATION AND DEPRESSION IN AD

4.1 Introduction

Depression in AD & the serotonergic system

Depression in Alzheimer's disease (AD) occurs in up to 50% of AD patients. It appears mainly in mild to moderate dementia, leads to faster cognitive decline and is associated with poorer quality of life and early institutionalization (Starkstein et al. 2008). The aetiology and pathophysiology of depressive symptoms in AD remains elusive. However, it has been hypothesised that they are the consequence of a distinct underlying genetic and neurobiological liability unmasked by the disease process rather than the results of a varying pattern of neurodegeneration in particular individuals (McIlroy & Craig 2004).

Pathophysiology of mood disorders and depression is thought to arise from a dysfunction in central nervous system (CNS) serotonergic circuits. The neurotransmitter serotonin (5-hydroxytriptamine, 5-HT) plays an acknowledged role in cognition, memory and regulation of sleep and mood. The raphe nuclei, which contain the majority of serotonin neurones, give rise to serotonergic projections that are widely distributed throughout the CNS and notably in brain areas critical for cognitive functions such as frontal cortex, temporal cortex and the hippocampal formation (Vertes et al. 1999). The effects of 5-HT occur both directly and indirectly through activation of 5-HT-specific receptors and the modulation of other neurotransmission systems including cholinergic, glutamatergic, dopaminergic and GABAergic (Buhot 1997). Support for the role of the serotonergic system in depression mainly comes from the observations of decreasing concentrations of 5-HT in cerebrospinal fluid (CSF) of depressed patients and the efficacy of drugs that increase serotonergic neurotransmission in depression (Ressler & Nemeroff 1999). Antidepressant medications, such as the selective serotonin reuptake inhibitors (SSRIs) and tricyclic antidepressants, directly inhibit the serotonin transporter (SERT) (Lotrich et al. 2008; Roman et al. 2003).

The serotonergic system is extensively affected in AD, as demonstrated by post-mortem studies of patients with AD pathology, which show alterations in various serotonergic markers, ranging from, a reduction in serotonergic neurons in the raphe nuclei and associated loss of cortical 5-HT projections (Chen et al. 2000; Hendricksen et al. 2004) to decreased 5-HT receptors in a

number of brain regions including the frontal, parietal and temporal cortices (Lai et al. 2002; Lai et al. 2005; Bowen et al. 2008). In addition Positron Emission Tomography (PET) studies demonstrate reduced levels of serotonin receptors in the hippocampus (Kepe et al. 2006; Truchot et al. 2007). AD related deficits in serotonin neurotransmission are associated with accelerated cognitive decline (Lai et al. 2002), behavioural symptoms (Garcia-Alloza et al. 2005) and the severity of dementia. Thus the serotonergic system may be responsible for the cognitive and non-cognitive defects associated with AD and treatments with SSRIs have been shown to increase the CSF concentration of 5-HT and improve cognitive function and memory in these patients (Mossello et al. 2008; Mowla et al. 2007).

Candidate genes for depression in AD are expected to be genes associated with mood disorders, such as the serotonin transporter gene, SLC6A4 (SERT). The SERT gene is the most examined candidate in depression and is a major determinant of serotonin reuptake and inactivation following release at synapses. Transcriptional activity of the SERT gene is regulated by several positive and negative regulatory elements within the SERT promoter region (Heils et al. 1996; Lesch et al. 1996).

SERT gene promoter polymorphism (5HTTLPR)

A polymorphism located ~1.2kb upstream of the transcription start site in the promoter region of the SERT gene (5HTTLPR) consists of two common alleles in European populations, a short (S) variant with 14 repeats of a variably conserved 20-23 bp units and a long (L) variant, which has 16 repeats (Lesch et al. 1996; Nakamura et al. 2000; Avula et al. 2011).

Transfection studies with expression vectors containing the L or S variants linked to a reporter gene indicate that the L variant directs higher levels of transcription compared to the S variant when expressed in various cell lines including the raphe nucleus-derived cell line RN46 (Lesch et al. 1996; Heils et al. 1996; Mortensen et al. 1999). In agreement with the reporter gene assays, lower SERT mRNA levels and lower rates of serotonin uptake were measured in lymphoblasts of individuals who had at least one S allele compared to L/L homozygotes additionally the S allele was associated with anxiety-related personality traits (Lesch et al. 1996). Whereby the biological mechanism, through which the 5HTTLPR S allele can affect depression, is thought to be mediated through a decrease in SERT mRNA expression, resulting in abnormal 5-HT homeostasis and transmission, which in turn could account for the manifestation of behavioural symptoms (Lesch et al. 1996; Heils et al. 1996). However a smaller study using an allelic

expression imbalance approach found no correlation between 5HTTLPR genotype and allelic expression ratios or SERT mRNA levels in B-lymphocytes (Lim et al. 2006).

Also whether these changes translate into alterations in functional protein levels has been the subject of much investigation by *in vivo* imaging studies (Heinz et al. 2000; Willeit et al. 2001; Shioe et al. 2003; Parsey et al. 2006). Although results from one of these *in vivo* studies (Heinz et al. 2000) was consistent with those in cell lines (Lesch et al. 1996; Heils et al. 1996; Mortensen et al. 1999), where the L/L homozygote individuals demonstrated greater SERT binding compared to the S allele carriers, others found no differences in observed binding to SERT between genotypes (Willeit et al. 2001; Shioe et al. 2003). Furthermore Parsey et al. (2006) found no effect of 5-HTTLPR genotype on serotonin transporter binding potential in medication-free patients with major depressive disorder or in healthy volunteers using a triallelic genotype classification, accounting for a further variant within the L allele (Parsey et al. 2006). Though the SERT S allele has been has been widely implicated in affective disorders, results of these studies have been mixed and although they have pinpointed an increased vulnerability to depression in carriers of the S allele, meta-analyses have not demonstrated a significant role for pure genetic effects of the polymorphism in depression (Lotrich and Pollock 2004; Lasky-Su et al. 2005).

SERT gene & AD

With regards to AD significant decreases in SERT have been observed in the temporal cortex, frontal cortex and the hippocampus (Tejani-Butt et al. 1995; Chen et al. 1996; Thomas et al. 2006; Lai et al. 2011). While imaging studies have confirmed the reductions in SERT binding at early stages of the disease in the dorsolateral prefrontal cortex, an area involved in emotional processing, cognition and executive function which is consistently compromised in depression (Ouchi et al. 2009). Others have suggested differential involvement of hippocampal serotonin 1A receptors and SERT sites, based on the observed loss of serotonin 1A receptors specifically correlated with depressive symptoms, while SERT sites were preserved or up-regulated in patients with aggressive behaviours in AD (Lai et al. 2011).

The relatively high frequency of the 5HTTLPR S variant has been suggested to be associated with an increased risk of AD in the Caucasian population (Hu et al. 2000). Studies in both the UK and Brazilian populations (Li et al. 1997; Oliveira et al. 1999) found the 5HTTLPR S allele to be elevated in AD subjects compared to controls and unrelated to the ε4 allele of the ApoE gene. A recent study in the Caucasian population has, however, failed to find any association between

polymorphism of the SERT promoter and increased risk of AD (Seripa et al. 2008), as previously found in similar analysis of Columbian and Japanese populations (Forero et al. 2006; Kunugi et al. 2000).

Some studies have linked the 5HTTLPR polymorphism to AD related behavioural symptoms including anxiety, psychosis and aggression. With some studies demonstrating the L allele to be associated with increased risk for anxiety or aggression and psychosis (Tsang et al. 2003; Quaranta et al. 2009; Sukonick et al, 2001; Sweet et al. 2001), while the S allele was associated with increased risk of psychosis in one study (Borroni et al. 2006). Conversely, similar analysis of the 5HTTLPR polymorphism in relation to depression in AD found there to be no significant effect of genotype on depression (Li et al. 1997; Assal et al. 2004; Micheli et al. 2006; Pritchard et al. 2007; Ueki et al. 2007; Grünblatt et al., 2009). Though the serotonergic system is affected in AD and the changes already present at very early stages of the disease, dysfunction in this resilient system may require further environmental stressors to develop depression.

SERT gene-environment

Emerging evidence has described gene-environment interactions between stressful life events (SLE) and the 5HTTLPR S allele modifying the risk for depression in children, adolescents, and adults. More specifically, the S allele was found to be associated with the development of depression, but only in individuals with a history of childhood maltreatment or recent SLE (Caspi et al. 2003). The authors reported that the number of SLEs positively correlated with subjects' self-report of depressive symptoms, and moreover, that individuals with the S/S or S/L genotype reported more severe depressive symptoms in response to SLEs, while individuals exposed to childhood maltreatment who possessed the S/S genotype had the highest probability of developing a major depressive episode, followed by the S/L genotype (Caspi et al. 2003). In addition, it has been reported that the 5HTTLPR polymorphism appears to influence the response to treatment with antidepressants (Lotrich et al. 2008; Rausch et al. 2002) and antipsychotics (Vazquez-Bourgon et al. 2010) and that gender can modify this interaction (Huezo-Diaz et al. 2009).

Early life stress (ELS) is also thought to influence neurodevelopment and subsequent behavioural response to stress. McFarlane et al. (2005) measured electrophysiological brain activity in non-clinical subjects, demonstrating that ELS has significant effects on brain function and the personality dimensions neuroticism and openness, while the number of ELS events experienced was shown to be a significant predictor of scores on the Depression Anxiety Stress

Scales questionnaire, which rates subjects on symptoms of depression. While depressed patients who both carried the 5HTTLPR S allele and had a positive history for early emotional neglect were noted to develop smaller hippocampal volumes compared to those with only one of these risk factors (Frodl et al. 2010). Raising the possibility that 5HTTLPR S allele carriers show an enhanced stress response and increased neuroplastic changes due to ELS that might affect neuronal development and in turn, may then increase the risk to depression (Frodl et al. 2010). Such ELS changes are further compounded by later life stress, in mice showing an interaction between adverse early life experience and stress during adulthood leading to increased slc6a4 mRNA expression within specific subdivisions of the dorsal raphe nucleus (Gardner et al. 2009).

Moreover a number of animal studies suggest it is possible the response to stressors maybe mediated through epigenetic changes (Champagne & Mashoodh 2009; Johnstone & Baylin 2010; Meaney 2010). For instance elevated stress sensitivity in mice that were maternally separated, was associated with increased hypothalamic expression of arginine vasopressin (AVP), a neuropeptide involved in the regulation of complex social cognition and behaviours along with decreased levels of DNA methylation at particular CpG dinucleotide within an AVP enhancer region (Murgatroyd et al. 2009).

SERT gene DNA methylation

The SERT gene is subjected to DNA methylation and recent studies have demonstrated that DNA methylation of a CpG island in the SERT promoter shows a trend for an association of increased overall methylation with lifetime history of depression (Philibert et al. 2008) and that depressive symptoms were more common among those with elevated SERT promoter methylation who carried the 5HTTLPR S allele (Olsson et al. 2010). While increased amounts of methylation at a number of individual CpG sites within the SERT promoter were significantly associated with decreasing mRNA, these did not withstand correction for multiple testing (Philibert et al. 2008). Though increased methylation at the SERT promoter has been shown to be associated with childhood abuse and related to symptoms of antisocial personality disorder (Beach et al. 2010, 2011), another study assessing the affects of stress on SERT promoter CpG methylation among nurses from high and low work stress environments, showed nurses in the high stress environment had significantly lower promoter methylation levels compared to nurses in the low stress environment, although there was no significant interaction of 5-HTTLPR genotype and work stress with methylation (Alasaari et al. 2012).

Both a cell line encoding a fluorescent protein and a luciferase reporter plasmid construct have been used to demonstrate a functional relationship between level of SERT promoter methylation and SERT mRNA expression indicating that even partial methylation can silence the promoter (Olsson et al 2010; Wang et al. 2012). More recently lower *in vivo* brain serotonin synthesis measured by PET was associated with higher levels of DNA methylation at certain CpG sites of the SERT promoter methylation in T cells and monocytes within a sample of adult males with high childhood-limited aggression (Wang et al. 2012).

Though very little is known about the role of the SERT CGI in illness, the amount of variance in mRNA and protein production accounted for by the 5HTTLPR is relatively modest (<10%) (Bradley et al. 2005), therefore it is likely that the CGI plays a regulatory role in SERT expression. In a maternal/social separation, rhesus macaque model of ELS the low expressing rh5HTTLPR S alleles exhibited higher mean SERT CpG methylation in blood, which was associated with lower SERT expression, however higher SERT CpG methylation, but not rh5-HTTLPR genotype, exacerbated the effects of early life stress on behavioural stress reactivity in infants (Kinnally et al. 2010). Though these data are based on blood samples they indicate that SERT CGI methylation is functionally related to SERT expression early in life and that methylation status of the CGI may explain some of the apparent relationship of stress and 5HTTLPR genotype with subsequent psychopathology. In contrast to genotype data, expression and methylation levels are dynamic, and can be affected by variables such as the ageing process, gender, treatment with medication and cognitive impairment. It is therefore very challenging to combine all this information on these different levels in order to get a more comprehensive picture of gene regulation and how this is associated with behavioural changes.

Both AD and depression are genetically complex disorders associated with multiple genetic defects and environmental factors contributing to their pathology and phenotypic expression. Though a number of studies to date have sought for associations between the 5HTTLPR polymorphism and depression in AD, results have provided no support for an association (Li et al. 1997; Assal et al. 2004; Micheli et al. 2006; Pritchard et al. 2007; Ueki et al. 2007; Grünblatt et al., 2009). However, some studies used different scales to assess behaviour and show inconsistencies on the adjustment of variables, which could confound these associations, with others not adjusting for any confounders. Additionally, no studies to date have examined gene X environment (such as SLE) interactions along with SERT methylation and mRNA levels in relation to depression in AD.

This chapter will initially explore the association between SERT promoter DNA methylation and mRNA abundance across seven different brain areas in relation to AD status and additionally look at the associations of 5HTTLPR genotype, the presence of SLEs and depression in AD samples as assessed by the Cornell Scale for Depression in Dementia (CSDD) (Alexopoulos et al. 1988). Furthermore clinical data including duration of the disease, age and sex of participants, cognitive impairment, years in education and treatment with antidepressants and antipsychotics, which have all been associated with depression in AD (Modrego 2010; Verkaik et al. 2007; Gilley et al. 2004), will be incorporated where applicable. In combing these data to understand whether SERT mRNA expression changes in the different brain areas associate with depression in AD, we should be able to test whether or not these changes are possibly modulated by 5HTTLPR polymorphic variation, environmental stresses, and DNA methylation at the promoter (Figure 4.1) as mentioned above.

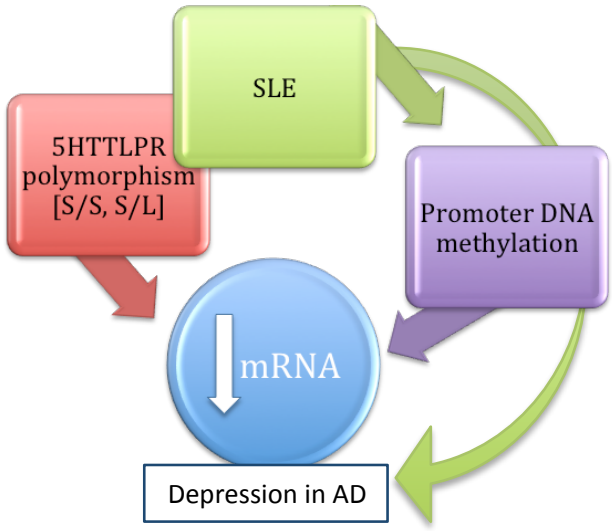


Figure 4.1 Environmental, genetic and epigenetic modulators of SERT expression & depression in AD

4.2 Aims

I. To measure DNA methylation at the SERT promoter in cases and controls across seven different brain regions of post-mortem tissue and to examine whether there are any differences in methylation levels that associate with disease status.

II. To examine the relationship of SERT promoter DNA methylation and mRNA expression levels in these samples.

III. To explore the relationship of the 5HTTLPR S allele and the presence of a SLE on depression in AD, along with their effect on SERT promoter methylation and mRNA expression.

4.3 Samples

All samples were taken from a well characterised cohort of brains donated to the Medical Research Council London Neurodegenerative Diseases Brain Bank. These consisted of both AD brains and matched cognitively normal controls (N=61), of which 37 are confirmed AD patients. Diagnosis was autopsy confirmed by an experienced pathologist. AD brain samples were matched with normal controls on sex and age where possible, though this was limited by the small number of normal control brains available for analysis. The mean age at death was 86.3 years (Range: 69–89) for AD cases and 77.1 years (Range: 40–92) for controls. Twenty seven patients were female (73%) and 10 males (27%) while in the control group, 12 subjects were females (50%) and 12 subjects were males (50%). A full list of the samples used in this chapter along with demographic details is listed in Appendix A.

Frozen brain tissue was collected from Broadmann areas, B8, B9, B10 and B46 as well from the Entorhinal Cortex, the Superior Temporal Gyrus and the Visual Cortex (control area). The Broadmann areas from the frontal cortex are areas specifically associated with behaviour, while the Entorhinal Cortex and Superior Temporal Gyrus reflect AD pathology affected tissue. DNA was extracted using a standard phenol-chloroform extraction (section 2.1).

Of the 37 AD patients 32 of them had full behavioural data including the Cornell Scale for Depression in Dementia (Alexopoulos et al. 1988), which was recorded at baseline interview. The Cornell Scale was specifically developed to assess major depression in patients with dementia. It uses a comprehensive interviewing approach that derives information from both the patient and the informant in two semi-structured interviews assessing behaviour in the week preceding the interview. It comprises of 19 items measuring 5 clusters of symptoms: mood related signs, behavioural disturbance, physical signs, cyclic functions and ideational disturbance. Each item is rated for severity on a scale of 0-2 (0 = absent, 1 = mild or intermittent, 2 = severe) and the item scores are added. This sum of scores was used in the present analysis ranging from 0 to 24.

Gender, years in full time education (coded as: 0 = less than 9 years in education, 1 = 9-12 years in education, 2 = 13 and more years in education), treatment with antidepressants and antipsychotics, interview age (years), disease duration (years) and cognitive impairment status measured with the Mini Mental Examination Score (MMSE; 0 = severe cognitive impairment, 1 = moderate to sever cognitive impairment, 2 = mild to moderate cognitive impairment) were all recorded for the AD patients at baseline interview. Demographic data were used in order to

quantify stressful life events (SLE) including history of illness, chronic disability, death of spouse and lone accommodation and the absence (0) or presence (1) of SLE was used.

4.4 Statistical analysis

In general linear regression models were used to test the association of predictor/independent variables (disease status, demographic or clinical data) with outcome/dependent variables (DNA methylation, mRNA levels or depression score). As a prelude to multivariate regression modeling, univariate regressions were carried out for individual independent variables with the dependent variable to ascertain if these variables needed to be added as a covariate to the final multivariate regression model; if a independent variable was significantly associated (p-value = <0.05) with the dependent variable it was included as a covariate in all further analyses. Unless otherwise stated all analyses were controlled for variables such as sex, age and ApoE genotype. Mixed effects models were used to combine information from the different brain areas taking into account both between patient and within patient variation, using the data from the visual cortex brain area as baseline. The brain area data presented in the mixed model outputs merely summarise if the given brain area is different from the VC baseline area in that model. For all models, regression coefficients (β) represent the change in the dependent variable (such as DNA methylation) for each unit change in the predictor/independent variable.

4.5 SERT (SLC6A4) promoter methylation content assayed

The SLC6A4 promoter region is characterised by the presence of both a CGI and a SERT 5HTTLPR polymorphism that is around 1.2 kb upstream of the Transcription Start Site (TTS). The whole CGI was covered by two bisulphite PCR amplicons, assayed on the Sequenom MALDI-TOF platform (Figure 4.2). CpG dinucleotide methylation content was assayed at the given CpG units as shown in Figure 4.4 and which correspond to those within the sequence of Figure 4.3 above it, note that not all the CpGs within the amplified sequences provide methylation information and hence the data is represented for the given MALDI-TOF CpG fragments. The PCR amplicons were designed and run on the Sequenom platform as described in section 2.9.

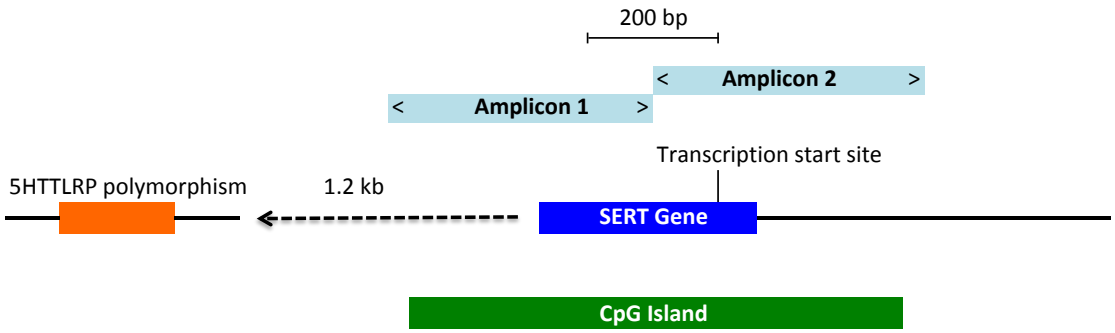


Figure 4.2 Schematic of the SLC6A4 promoter region

TAtTGtTAGGTtttAGGAAGAAAGAGAGAGtAGtTTT*cgGGATGGGGAcgATGGGGAGGT
GTtcgAGGTtAAGAGAAAGcgGtAcgAGtAGAttttTGTGTGtcgTttTGTGGGcgcgGGG
cgGtAGGGGAGGcgtAtAttTGtTtttTTTGTGtAGttTttttttTttcgtAAAGTTAAAGA
GtAGGAAAGTtAGGATTttTcgtTcgGtttTGtttTGtcgGtTGtTtcgcgtTtcgtTttT
tttTGcgAGcgTGTGTGTGTGTcgGGGGTtttTttttTttTGGtTtTGGGGTcgGGcgcgt
AtttcgtttcgtagcgcggttttttttTGGcgAGcgtAAttttATttAGcgGGAGcgcgGA
GtcgcgGtcgcgGGGAAGtATTAAGTTTATTcgttTtAAAGTGAcgtAAAAATTtTTtAAG
AGtTtTTTGGcgGcgGtTATtTAGAGATtAGAttATGTGAGGGttcgcgGGTAtAAATAcg
GtcgcgtcgcgtttttttcgtAtAGttAGcgtcgtcgGGTGttTcgAGGGcgcgAGGttAG
ttcgttTGtttAGttcgGGAttAGttTttcgcgtAGttTGGtAGGTGGGTtcgtTTTTtt
TtTtcgttTcgAAtttAcgTTTtTTTttAGAtttTTtTTttcgttTcgGGGAGGGGGATAG
AAtcgtTGcgttttAtcgtttTGcgAGGAGGcgAGGAGGTGtATGcgttttAGcgGTGGGc
gtcgGATttTGttttTGcgttTTtAcgtTtAGtAAGAGttAGAGtTGAAGtTGAtcgGtt
AGAGTGGGAGAcgAGGAAcgTGGAGTGtTcgAAGTGGGcgGGcg*TAGGGGGtTttTTTGT
tTATTGTTGtAGG

Figure 4.3 Bisulphite converted sequence of the SLC6A4 CGI
[*: start and end of CGI; grey font: primer sequences, N.B. amplicon 1 Rev and amplicon 2 Fwd cover the same sequence; underlined: exon 1; lower case t: bisulphite converted C; bold lower case cg: assayed CpG; red font: SNPs; blue font: CpG units corresponding to amplicon 2 units b22, b30 and b40

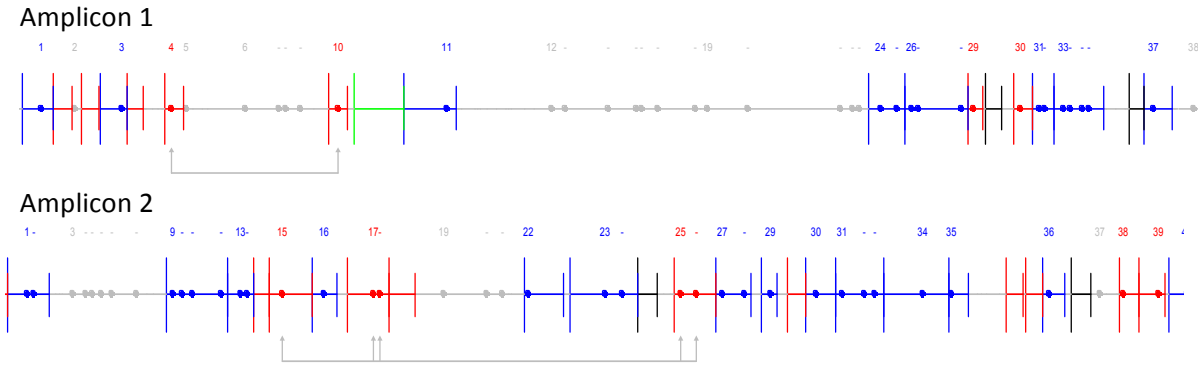


Figure 4.4 Predicted assayed MALDI-TOF CpG units of amplicon 1 and 2
[blue: assayable fragment with CpGs; red: assayable fragment with CpGs and a MW overlap; grey: unassayable fragment with CpGs]

4.6 Results

4.6.1 Case-control analysis of SERT promoter CGI methylation

Initial analysis was carried out, looking at the whole cohort of samples, both Alzheimer's and controls (N=61) for the seven different brain regions. Average methylation levels across the whole CGI are in the range of 0.063-0.160% with the VC control brain region having the lowest mean value of 0.099% (Figure 4.5). Data was generated across 13 CpG assayed fragments in amplicon 1 and 18 in amplicon 2. Methylation levels across amplicon 1 (range 0.015-0.122%) are lower than those obtained for amplicon 2, which crosses over into the transcribed sequence (range 0.084-0.255%) (Figure 4.8). Overall these data demonstrate that the CGI has low levels of DNA methylation.

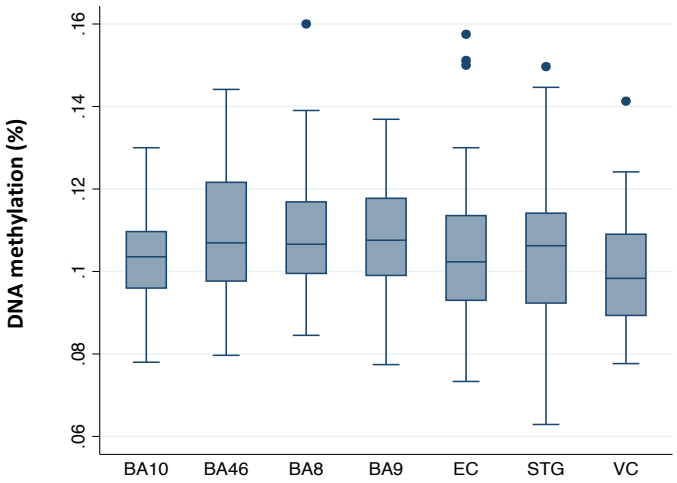


Figure 4.5 Average methylation results at the SERT promoter across all seven different brain regions [BA: Brodmann Area; EC: Entorhinal Cortex; STG: Superior Temporal Gyrus; VC: Visual Cortex; whiskers: 95% confidence interval]

Exploratory Analysis

The observed methylation data in Figure 4.5 show that there are extremely low levels of methylation at the SERT CGI, more importantly these are below the technical detection level of the Sequenom MassARRAY EpiTYPER platform, which is around 5% (Coolen et al. 2007). Given that all the data are below 5% and that any measurable differences are very small and not likely to be significant, the remainder of this chapter will be purely exploratory. Where associations are noted as being significant these are not likely to be robust enough to withstand corrections for multiple testing and are thus only presented as exploratory findings.

In general the AD samples show a trend for increased average CGI methylation compared to controls across all brain regions except at the BA9 area (Figure 4.6). Using a linear regression model of average CGI methylation levels, for each brain area in turn, only the BA10 brain region showed status to be significantly associated with average CGI methylation levels (β = 0.009, SE = 0.004, *p* = 0.024). A mixed model using the VC as a baseline was used to further inspect the association of status with average CGI methylation levels at the BA10 brain region and to see if there are significant differences between this behaviour associated BA10 brain area and control brain region VC. While status remained marginally significantly associated with average CGI methylation (β = 0.006, SE = 0.003, *p* = 0.046) no significant differences were observed between average CGI methylation levels at the BA10 and VC brain regions.

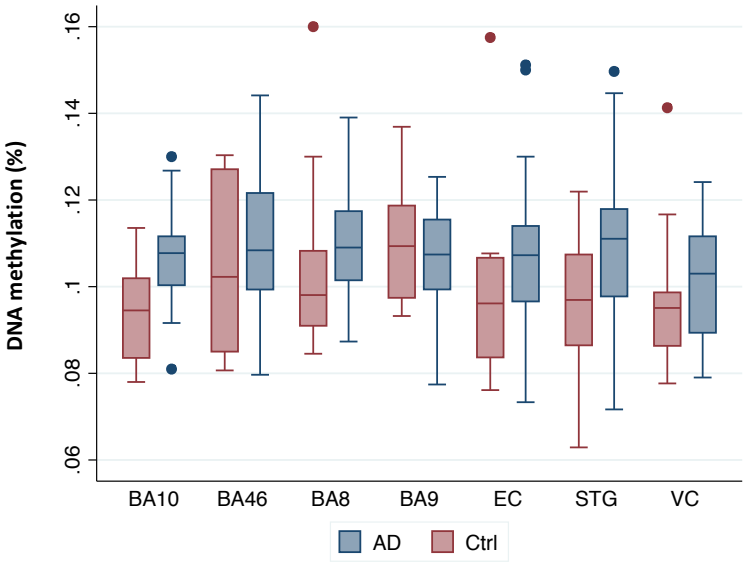


Figure 4.6 Average methylation at the SERT CGI across brain regions for AD and controls [BA: Brodmann Area; EC: Entorhinal Cortex; STG: Superior Temporal Gyrus; VC: Visual Cortex; AD: Alzheimer's disease, n=37; Ctrl: controls, n=24; whiskers: 95% confidence interval]

In an attempt to combine the data from all brain regions and identify any significant differences or associations a mixed model of average CGI methylation levels was used. While AD status was not significantly associated with average CGI methylation, frontal cortex regions BA9, BA46 and BA8 had significantly higher levels of average CGI methylation (BA9: β = 0.007, SE = 0.003, *p* = 0.011; BA46: β = 0.008, SE = 0.003, *p* = 0.01; and BA8: β = 0.008, SE = 0.003, *p* = 0.009) compared to the baseline VC area (Table 4.1). Moreover there is an overall significant association of age at death with average CGI methylation (β = 0.0004, SE = 0.0001, *p* = <0.001) as depicted in Figure 4.7, with a 0.0004% increase in average SERT CGI methylation with each year increase in age. While these data are not from the same individuals across different time points, they indicate a relationship of SERT CGI promoter methylation with time.

Variable		Coefficient	Std. Err.	Z	*p*	95% CI	
AD status		0.0018	0.0025	0.7200	0.4710	-0.0031	0.0068
Sex		-0.0042	0.0023	-1.8100	0.0700	-0.0088	0.0003
*Age at death		0.0004	0.0001	3.6900	<0.0001	0.0002	0.0007
ApoE genotype		-0.0001	0.0002	-0.5300	0.5950	-0.0005	0.0003
Brain region:	*BA9	0.0071	0.0028	2.5400	0.0110	0.0016	0.0126
	BA10	0.0032	0.0029	1.1300	0.2590	-0.0024	0.0089
	*BA46	0.0078	0.0030	2.5600	0.0100	0.0018	0.0137
	*BA8	0.0077	0.0030	2.6200	0.0090	0.0019	0.0135
	EC	0.0046	0.0032	1.4500	0.1480	-0.0016	0.0109
	STG	0.0044	0.0031	1.4200	0.1560	-0.0017	0.0104

Table 4.1 Mixed model output of variables on average CGI methylation levels [BA: Brodmann Area; EC: Entorhinal Cortex; STG: Superior Temporal Gyrus; VC: Visual Cortex; AD: Alzheimer's disease; *: significant variable]

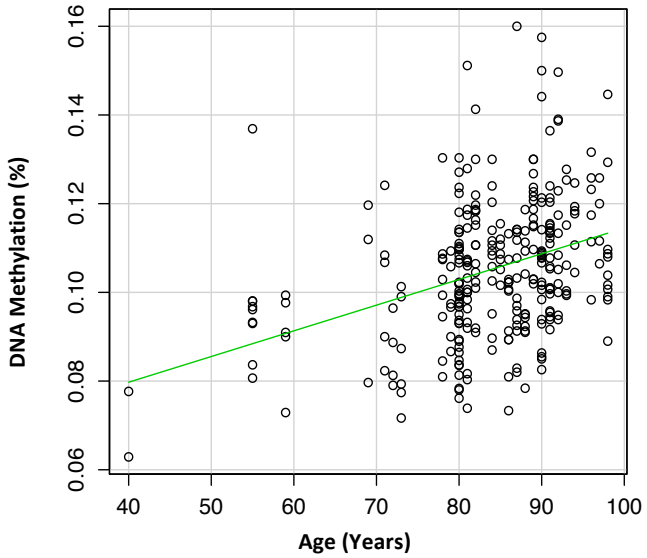


Figure 4.7 Scatterplot of average SERT CGI methylation levels by age at death [green line: line of least squares]

With regards to a case-control comparison of average CGI methylation a number of CpG fragments, particularly in amplicon 2 show a slight increase in methylation for the AD samples (Figure 4.8). Given these observed differences methylation data from the two amplicons was then treated separately. In separate mixed models for the average methylation levels at the CGI for amplicon 1 and amplicon 2, methylation at amplicon 1 showed no significant associations or differences by brain area in methylation levels (Table 4.2). Whereas methylation at amplicon 2 presented a significant association of age and sex with average methylation (β = 0.0008, SE = 0.0002, *p* = <0.001 and β = -0.007, SE = 0.004, *p* = 0.047 respectively), where males have less methylation at the SERT CGI amplicon 2 than females (Figure 4.9). Also there are significant differences in average methylation at regions BA9, BA46 and STG (BA9: β = 0.013, SE = 0.005, *p* = 0.007; BA46: β = 0.012, SE = 0.005, *p* = 0.015; and STG: β = 0.014, SE = 0.005, *p* = 0.003) compared to the baseline VC area (Table 4.3).

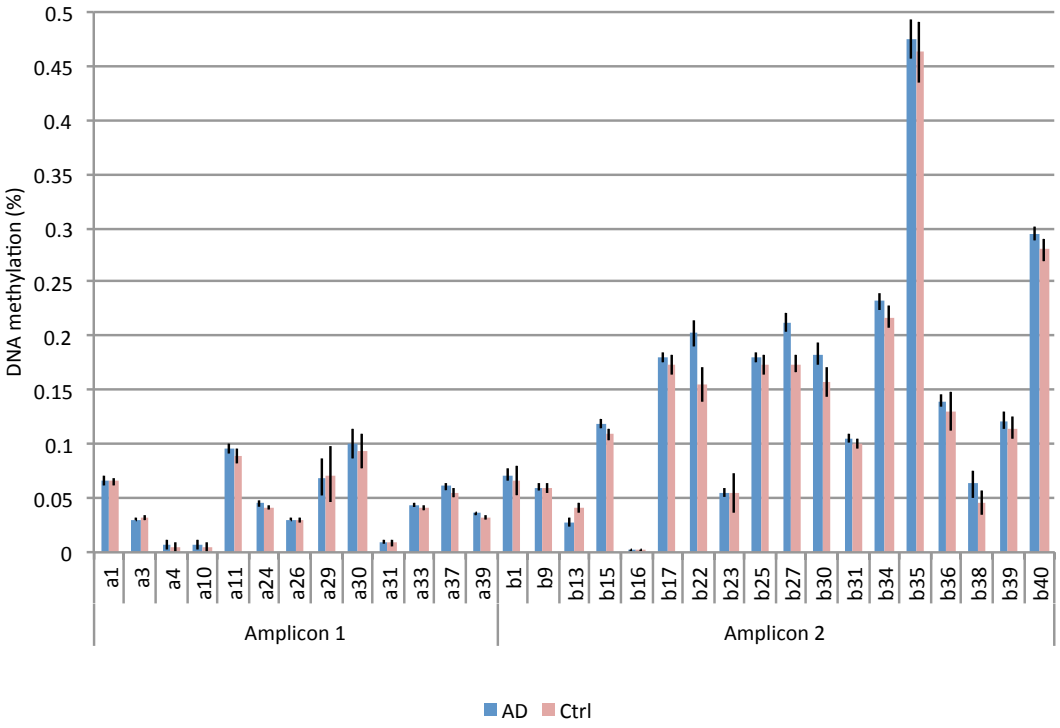


Figure 4.8 Average methylation measured across the brain at each assayable CpG unit by disease status [AD: Alzheimer's disease, n=37; Ctrl: controls, n=24; whiskers: 95% confidence interval]

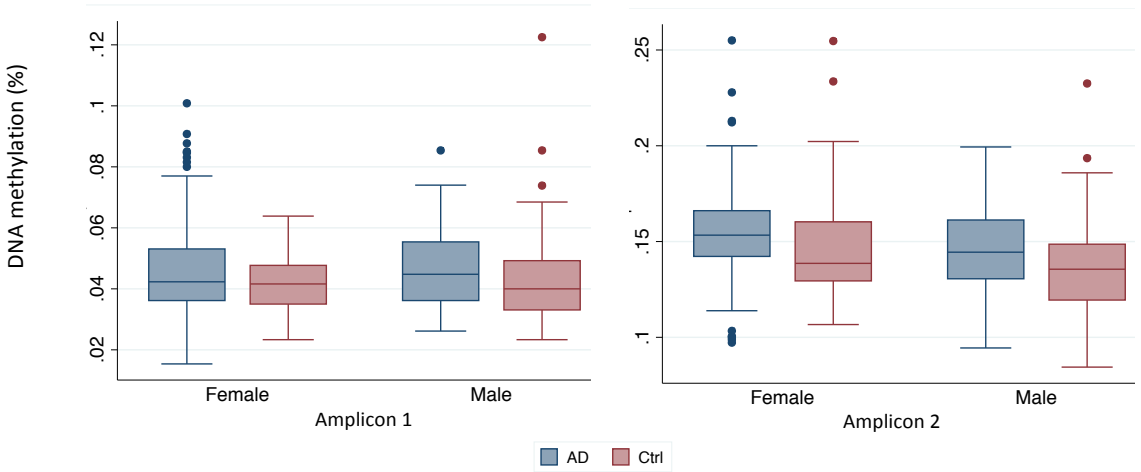


Figure 4.9 Average methylation at the SERT CGI amplicons 1 & 2 by sex and disease status [; AD: Alzheimer's disease, female n=27, male n=10; Ctrl: controls, female n=12, male n=12; whiskers: 95% confidence interval]

Variable		Coefficient	Std. Err.	Z	*p*	95% CI	
AD status		0.0020	0.0021	0.9600	0.3370	-0.0021	0.0061
Sex		0.0014	0.0019	0.7200	0.4690	-0.0023	0.0051
Age at death		0.0001	0.0001	1.3600	0.1720	-0.0001	0.0003
ApoE genotype		0.0000	0.0001	0.1300	0.8970	-0.0003	0.0003
Brain region:	BA9	0.0021	0.0026	0.8100	0.4190	-0.0030	0.0072
	BA10	0.0012	0.0027	0.4600	0.6490	-0.0040	0.0065
	BA46	0.0026	0.0028	0.9300	0.3500	-0.0029	0.0082
	BA8	0.0048	0.0028	1.7100	0.0870	-0.0007	0.0102
	EC	0.0003	0.0029	0.1000	0.9210	-0.0055	0.0060
	STG	-0.0003	0.0029	-0.0900	0.9310	-0.0059	0.0054

Table 4.2 Mixed model output of variables including status on average amplicon 1 CGI methylation levels [BA: Brodmann Area; EC: Entorhinal Cortex; STG: Superior Temporal Gyrus; VC: Visual Cortex; AD: Alzheimer's disease; *: significant variable]

Variable		Coefficient	Std. Err.	Z	*p*	95% CI	
AD status		0.0013	0.0038	0.3400	0.7320	-0.0061	0.0087
*Sex		-0.0069	0.0035	-1.9900	0.0470	-0.0138	-0.0001
*Age at death		0.0008	0.0002	4.2900	<0.0001	0.0004	0.0011
ApoE genotype		-0.0001	0.0003	-0.2700	0.7850	-0.0006	0.0005
Brain region:	*BA9	0.0127	0.0047	2.7100	0.0070	0.0035	0.0219
	BA10	0.0078	0.0048	1.6400	0.1010	-0.0015	0.0172
	*BA46	0.0122	0.0050	2.4300	0.0150	0.0024	0.0220
	BA8	0.0092	0.0048	1.9000	0.0570	-0.0003	0.0186
	EC	0.0086	0.0051	1.6700	0.0940	-0.0015	0.0186
	*STG	0.0139	0.0048	2.9200	0.0030	0.0046	0.0233

Table 4.3 Mixed model output of variables including status on average amplicon 2 CGI methylation levels [BA: Brodmann Area; EC: Entorhinal Cortex; STG: Superior Temporal Gyrus; VC: Visual Cortex; AD: Alzheimer's disease; *: significant variable]

Also at the individual CpG fragment methylation levels the only CpG fragment to show disease status as being significantly associated with methylation levels was that seen for a association with average methylation levels at fragment b22 (β = 0.040, SE = 0.015, *p* = 0.007), with the BA46 and the STG regions exhibiting significantly higher methylation levels compared to the VC area (β = 0.036, SE = 0.017, *p* = 0.033; β = 0.035, SE = 0.017, *p* = 0.041 respectively) (Figure 4.8). In adopting Braak staging instead of AD status in all the above analyses, though a positive relationship was observed between increasing Braak stage with increasing levels of methylation none of these were significant.

4.6.1.1 Case-control analysis of SERT CGI methylation and mRNA expression levels

SERT mRNA levels were available for all the corresponding brain samples. As shown in Figure 4.10 the highest mRNA levels were observed for BA9 whereas the lowest were seen in the EC and the STG. The rest of the three frontal cortex areas exhibited expression levels similar to the VC.

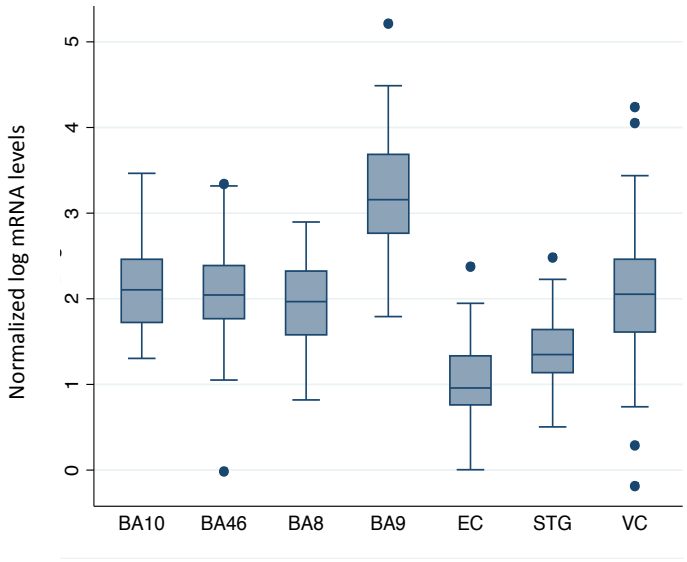


Figure 4.10 Average mRNA levels across all seven different brain regions
[BA: Brodmann Area; EC: Entorhinal Cortex; STG: Superior Temporal Gyrus; VC: Visual Cortex; whiskers: 95% confidence interval]

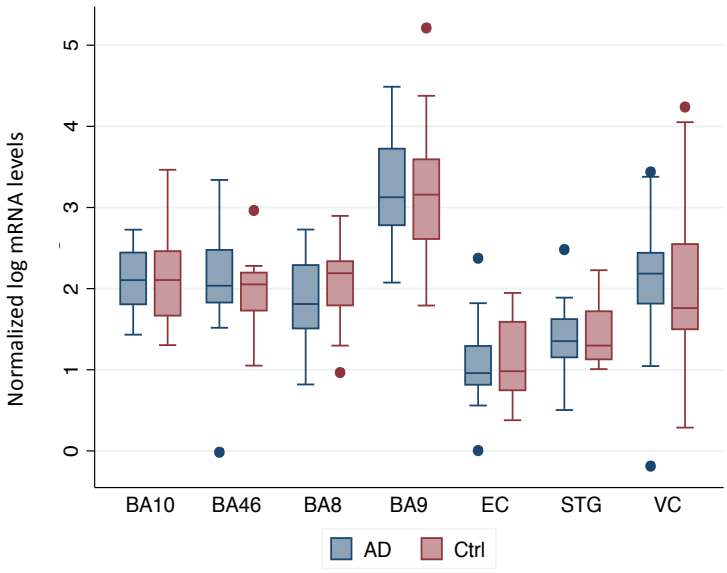
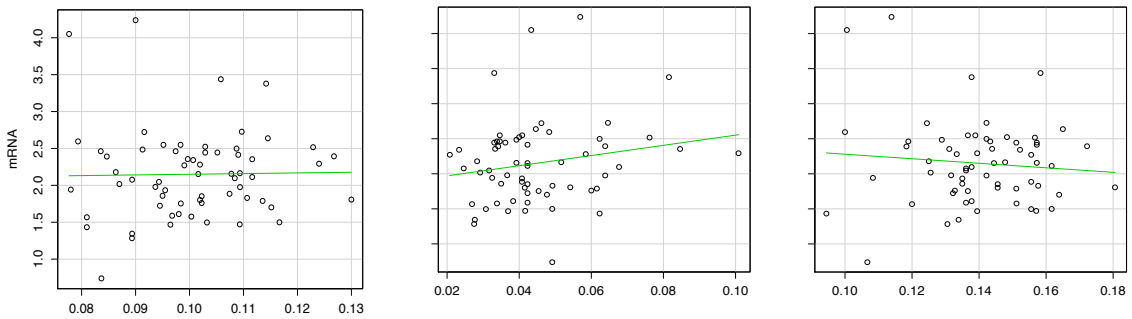


Figure 4.11 Average SERT mRNA levels across brain regions for AD and controls
[BA: Brodmann Area; EC: Entorhinal Cortex; STG: Superior Temporal Gyrus; VC: Visual Cortex; AD: Alzheimer's disease, n=37; Ctrl: controls, n=24; whiskers: 95% confidence interval]

In a simple attempt to look at the overall relationship of the different measures of SERT promoter methylation levels (average for the whole CGI, average at amplicon 1 and average at amplicon 2) against mRNA levels a scatterplot and Pearson correlation coefficient was calculated. There is no significant relationship between methylation levels and mRNA levels (Figure 4.12).



r = 0.02 , *p* = 0.88 *r* = 0.19 , *p* = 0.13 *r* = -0.09 , *p* = 0.46

Figure 4.12 Scatterplots & Pearson correlation coefficient of methylation measures against mRNA levels [green line: line of least squares]

Unlike the trend for increased average SERT CGI methylation in AD samples compared to controls seen in Figure 4.6, the SERT mRNA levels for AD and controls are roughly the same for each brain region, with the slight exceptions of regions BA8 and VC (Figure 4.11). In exploring the relationship between mRNA levels from the different brain regions and average CGI methylation levels in a mixed model of mRNA, disease status was not significantly associated with mRNA levels (β = -0.03, SE = 0.12, *p* = 0.81) (Table 4.5). While average CGI methylation was significantly associated with expression (β = 7.2, SE = 2.4, *p* = 0.003) albeit in a positive direction, along with significantly different mRNA levels at BA9, BA8, EC and STG compared to the VC (BA9: β = 0.99, SE = 0.12, *p* = <0.001; BA8: β = -0.36, SE = 0.13, *p* = 0.007; EC: β = -1.03, SE = 0.14, *p* = <0.001; and STG: β = -0.84, SE = 0.14, *p* = <0.001) (Table 4.5).

Variable		Coefficient	Std. Err.	Z	*p*	95% CI	
AD status		-0.029	0.121	-0.240	0.811	-0.266	0.209
*Average CGI methylation		7.149	2.432	2.940	0.003	2.382	11.916
Sex		0.173	0.118	1.470	0.142	-0.058	0.405
Age at death		-0.009	0.006	-1.520	0.129	-0.021	0.003
ApoE genotype		0.013	0.009	1.430	0.153	-0.005	0.030
Brain region:	*BA9	0.988	0.122	8.070	<0.001	0.748	1.228
	BA10	-0.154	0.127	-1.210	0.227	-0.403	0.096
	BA46	-0.167	0.139	-1.200	0.229	-0.439	0.105
	*BA8	-0.355	0.131	-2.700	0.007	-0.612	-0.097
	*EC	-1.030	0.147	-7.030	<0.001	-1.317	-0.743
	*STG	-0.839	0.143	-5.860	<0.001	-1.120	-0.558

Table 4.5 Mixed model output 1 of variables on SERT mRNA expression levels (average CGI methylation) [BA: Brodmann Area; EC: Entorhinal Cortex; STG: Superior Temporal Gyrus; AD: Alzheimer's disease; *: significant variable]

When separating out the average CGI methylation into CGI amplicon 1 (Table 4.6) and amplicon 2 (Table 4.7) in two additional mixed models, presented similar mRNA level differences at BA9, BA8, EC and STG compared to the VC. As might be expected only the association of average methylation at CGI amplicon 1 with mRNA levels was significant (β = 5.14, SE = 2.312, *p* = 0.026, Table 4.6), again in a positive direction as seen with the CGI average methylation levels (Table 4.5). Indicating that the methylation levels at amplicon 1, which is at the 5' end of the CGI and over the promoter region are possibly more related to SERT mRNA levels, while the methylation levels in the amplicon 2 over the exon 1 are less relevant to mRNA expression.

In addition the CpG fragment b22 identified in the section above to associate with AD status, showed no significant association with SERT mRNA levels when tested. Furthermore the CpG unit at fragment b34 has previously been reported to be associated with SERT mRNA expression levels (Philibert et al. 2008), though when tested within this dataset the association was in the expected direction, nonetheless it was not significant (β = -0.33, SE = 0.65, *p* = 0.61). In adopting Braak staging instead of AD status in all the above analyses, though a negative relationship was observed between increasing Braak stage with decreasing levels of mRNA none of these associations were significant.

Variable		Coefficient	Std. Err.	Z	*p*	95% CI	
AD status		-0.017	0.106	-0.160	0.876	-0.224	0.191
*Average CGI methylation amplicon 1		5.136	2.314	2.220	0.026	0.602	9.671
Sex		0.134	0.102	1.320	0.187	-0.065	0.334
Age at death		-0.010	0.005	-1.950	0.051	-0.020	0.000
ApoE genotype		0.014	0.008	1.810	0.070	-0.001	0.029
Brain region:	*BA9	1.090	0.110	9.870	<0.001	0.874	1.307
	BA10	-0.100	0.120	-0.830	0.404	-0.335	0.135
	BA46	-0.085	0.127	-0.670	0.503	-0.333	0.163
	*BA8	-0.276	0.123	-2.240	0.025	-0.517	-0.035
	*EC	-1.001	0.137	-7.330	<0.001	-1.269	-0.733
	*STG	-0.766	0.137	-5.570	<0.001	-1.035	-0.496

Table 4.6 Mixed model output 2 of variables on SERT mRNA expression levels (amplicon 1) [BA: Brodmann Area; EC: Entorhinal Cortex; STG: Superior Temporal Gyrus; AD: Alzheimer's disease; *: significant variable]

Variable		Coefficient	Std. Err.	Z	*p*	95% CI	
AD status		-0.056	0.118	-0.480	0.633	-0.288	0.175
Average CGI methylation amplicon 2		2.728	1.443	1.890	0.059	-0.100	5.555
Sex		0.191	0.116	1.650	0.099	-0.036	0.419
Age at death		-0.008	0.006	-1.290	0.197	-0.020	0.004
ApoE genotype		0.013	0.009	1.530	0.127	-0.004	0.031
Brain region:	*BA9	1.129	0.123	9.180	<0.001	0.888	1.370
	BA10	-0.069	0.129	-0.530	0.595	-0.321	0.184
	BA46	-0.040	0.141	-0.280	0.776	-0.315	0.235
	BA8	-0.207	0.133	-1.560	0.119	-0.468	0.053
	*EC	-0.945	0.149	-6.360	<0.001	-1.236	-0.654
	*STG	-0.791	0.144	-5.480	<0.001	-1.074	-0.508

Table 4.7 Mixed model output 3 of variables on SERT mRNA expression levels (amplicon 2) [BA: Brodmann Area; EC: Entorhinal Cortex; STG: Superior Temporal Gyrus; AD: Alzheimer's disease; *: significant variable]

4.6.2 Alzheimer's disease subset & clinical features

Of the 37 AD samples in this cohort 32 AD patients in this cohort were further characterized for the 5HTTLPR polymorphism and also had further clinical/behavioural measures such as, treatment with antidepressants and antipsychotics, cognitive impairment status measured with the Mini Mental Examination Score (MMSE), the presence of stressful life events (SLE) and depression scores as measured by the Cornell Scale for Depression in Dementia (CSDD) along with further characteristics which are displayed in Tables 4.8-4.9 below. Also largely the same patterns and levels of SERT mRNA and CGI methylation were observed across the different brain regions as those for the wider cohort described in the previous section.

With regards to the 5HTTLPR polymorphism initial observations indicated that the absence or presence of the S/S allele as opposed to the three possible genotypes provided a better model

fit in terms of SERT mRNA expression levels, hence this categorization was used for all the analyses that follow (Figure 4.13).

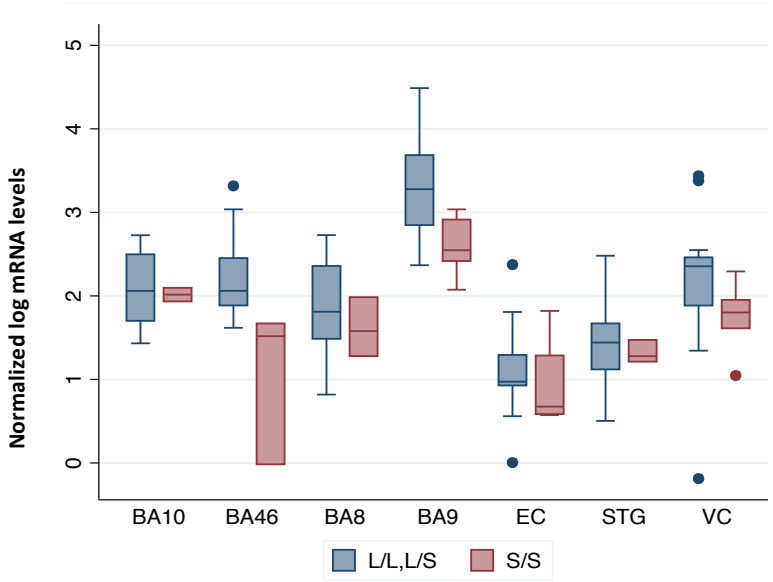


Figure 4.13 Average SERT mRNA levels and 5HTTLPR recessive model genotype by tissue [BA: Brodmann Area; EC: Entorhinal Cortex; STG: Superior Temporal Gyrus; VC: Visual Cortex; L/L,L/S: 5HTTLPR genotype, n=27; S/S: 5HTTLPR genotype, n=5; whiskers: 95% confidence interval]

Variable	Mean	Std. Dev.	Range
Age at death (years)	86.6	7.4	69-98
Age at score (years)	83.4	7.6	65-95
Disease duration (years)	9.1	3.7	0-19
Disease duration at score (years)	6.7	3.5	1.5-15
CSDD score	9.3	5.8	0-22
SERT mRNA levels (log)	2.05	0.86	-0.19-4.49
Average SERT CGI methylation (%)	0.11	0.01	0.07-0.15

Table 4.8 Characteristics of AD sample continuous variables [CSDD: Cornell Scale for Depression in Dementia]

Variable		Number	Percentage
Females/males		22/10	69/31
Presence/absence of SLE		25/7	78/22
Antipsychotics treated/untreated		8/24	25/75
Antidepressants treated/untreated		5/27	16/84
5HTTLPR	0 short S alleles	8	25
	1 short S allele	19	59
	2 short S alleles	5	16
MMSE	Severe impairment	15	46.9
	Moderate - severe impairment	10	31.25
	Mild - moderate impairment	7	21.9
Education	<9 years	16	50
	9-12 years	12	37.5
	>12 years	4	12.5

Table 4.9 Clinical and behavioural characteristics of AD sample categorical variables [SLE: Stressful Life Events; MMSE: Mini Mental Examination Score]

4.6.2.1 Modelling the effects of predictor variables on CSDD scores

Multiple linear regressions of CSDD scores with the variables outlined in Tables 4.8-4.9 demonstrate that the presence of a 5HTTLPR S/S genotype (β = -2.44, SE = 1.0, *p* =0.016) and the presence of a SLE (β = -3.03, SE = 0.89, *p* = 0.001) are significantly associated with CSDD scores, albeit in a negative direction. However testing for an interaction between the presence of a S/S genotype and a SLE, proved negative. Furthermore the disease duration and treatment with antidepressants were also significantly associated with CSDD scores (β = 0.02, SE = 0.008, *p* =0.019 and β = 6.7, SE = 1.1, *p* = <0.001 respectively) as one might expect (Table 4.10).

In relation to a much larger study of these clinical/behavioural measures (N=821) within which this dataset (N=32) is nested, a similar observation was made that the presence of each S 5HTTLPR allele was associated with a decrease in the mean CSDD score by 0.661 units (*p* = 0.004). However the observation that the presence of a SLE in this smaller sample set is also associated with decreasing CSDD scores is in the opposite direction to that observed in the larger study (β = 0.959, *p* = 0.008). Possibly indicating that this smaller sample set is underpowered or biased to detect the true effect of SLE on CSDD score. Given this the SLE association with CSDD scores is likely to be an unreliable estimate.

Variable	Coefficient	Std. Err.	t	*p*	95% CI	
Age at score	-0.028	0.054	-0.520	0.606	-0.134	0.078
*Sex	-1.829	0.867	-2.110	0.036	-3.538	-0.120
*ApoE genotype	-0.245	0.060	-4.070	<0.001	-0.364	-0.126
Antipsychotics	0.774	0.918	0.840	0.400	-1.035	2.583
*Antidepressants	6.734	1.115	6.040	0.000	4.537	8.932
MMSE category	-0.189	0.457	-0.410	0.679	-1.089	0.710
*Disease duration	0.020	0.008	2.350	0.019	0.003	0.037
*SLE	-3.033	0.888	-3.410	0.001	-4.783	-1.282
*5HTTLPR S/S genotype	-2.439	1.007	-2.420	0.016	-4.424	-0.454
Education category	-0.630	0.538	-1.170	0.242	-1.690	0.430

Table 4.10 Multiple linear regression output of variables on CSDD scores
[SLE: Stressful Life Events; MMSE: Mini Mental Examination Score]

4.6.2.2 Modelling the effects of predictor variables on SERT CGI methylation

To test the relationship of these clinical and behavioural measures on the SERT CGI methylation a mixed model with average CGI methylation was run (Table 4.11) with all the same fixed and possibly confounding variables indicated from the multiple linear regression above (Table 4.10). Both age at death (β = 0.0005, SE = 0.0002, *p* = 0.012) and the presence of SLE (β = 0.006, SE =

0.003, p = 0.038, differences depicted in Figure 4.14) were significantly associated with average CGI methylation, with the BA8 brain region showing significantly different levels of methylation compared to the VC (β = 0.007, SE = 0.004, p = 0.047). The presence of a 5HTTLPR S/S genotype, disease duration and treatment with antidepressants were not significantly associated with average CGI methylation levels (Table 4.11).

Variable		Coefficient	Std. Err.	Z	p	95% CI	
CSDD score		0.0005	0.0003	1.8600	0.0620	0.0000	0.0010
*SLE		0.0069	0.0034	2.0700	0.0380	0.0004	0.0135
5HTTLPR S/S genotype		-0.0011	0.0041	-0.2700	0.7870	-0.0092	0.0070
*Age at death		0.0005	0.0002	2.5000	0.0120	0.0001	0.0009
Sex		-0.0002	0.0034	-0.0700	0.9470	-0.0068	0.0064
ApoE genotype		0.0002	0.0002	1.1000	0.2720	-0.0002	0.0007
Antidepressants		-0.0050	0.0040	-1.2500	0.2100	-0.0128	0.0028
Disease duration		-0.0002	0.0004	-0.3900	0.6930	-0.0009	0.0006
Brain region:	BA9	0.0035	0.0031	1.1200	0.2650	-0.0026	0.0096
	BA10	0.0029	0.0032	0.8900	0.3750	-0.0035	0.0092
	BA46	0.0035	0.0035	0.9800	0.3280	-0.0035	0.0104
	*BA8	0.0074	0.0037	1.9800	0.0470	0.0001	0.0146
	EC	0.0057	0.0042	1.3800	0.1660	-0.0024	0.0139
	STG	0.0085	0.0046	1.8300	0.0670	-0.0006	0.0175

Table 4.11 Mixed model output of variables on average CGI methylation levels
[BA: Brodmann Area; EC: Entorhinal Cortex; STG: Superior Temporal Gyrus; CSDD: Cornell Scale for Depression in Dementia; SLE: presence of a stressful life event; *: significant variable]

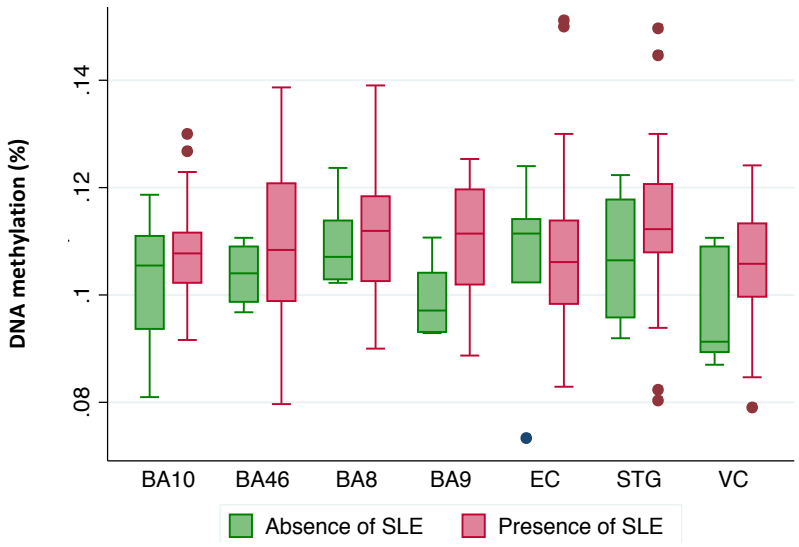


Figure 4.14 Average CGI methylation and the absence/presence of SLE by tissue
[BA: Brodmann Area; EC: Entorhinal Cortex; STG: Superior Temporal Gyrus; VC: Visual Cortex; Absence of SLE, n=7; Presence of SLE, n=25; whiskers: 95% confidence interval]

Given the differences observed in methylation levels at amplicon 1 and amplicon 2 the same model was applied to these separately. The amplicon 1 model (Table 4.12) shows no significant associations, while the BA8 brain region shows marginally significant different levels of methylation compared to the VC (β = 0.007, SE = 0.004, p = 0.047). As observed with the

average CGI methylation outcome model above (Table 4.11), the amplicon 2 model show (Table 4.13), both age at death (β = 0.0009, SE = 0.0003, *p* = 0.003) and the presence of a SLE (β = 0.010, SE = 0.004, *p* = 0.02) were significantly associated with average amplicon 2 CGI methylation, in addition the CSDD score was significantly associated with average amplicon 2 CGI methylation (β = 0.0008, SE = 0.0003, *p* = 0.015) with the STG brain region showing significantly different levels of methylation compared to the VC (β = 0.014, SE = 0.006, *p* = 0.019).

Variable		Coefficient	Std. Err.	Z	*p*	95% CI	
CSDD score		0.0002	0.0002	0.8000	0.4260	-0.0003	0.0007
SLE		-0.0008	0.0032	-0.2600	0.7980	-0.0071	0.0055
5HTTLPR S/S genotype		-0.0018	0.0038	-0.4600	0.6470	-0.0093	0.0058
Age at death		0.0003	0.0002	1.6400	0.1000	-0.0001	0.0007
Sex		0.0013	0.0032	0.4200	0.6770	-0.0049	0.0076
ApoE genotype		0.0000	0.0002	0.1300	0.8950	-0.0004	0.0004
Antidepressants		-0.0060	0.0038	-1.5800	0.1130	-0.0133	0.0014
Disease duration		-0.0003	0.0004	-0.7200	0.4740	-0.0010	0.0005
Brain region:	BA9	-0.0003	0.0035	-0.0800	0.9360	-0.0072	0.0066
	BA10	0.0040	0.0035	1.1500	0.2510	-0.0029	0.0109
	BA46	0.0051	0.0037	1.3700	0.1720	-0.0022	0.0123
	*BA8	0.0073	0.0037	1.9700	0.0480	0.0001	0.0146
	EC	0.0012	0.0039	0.3100	0.7570	-0.0065	0.0089
	STG	0.0014	0.0041	0.3300	0.7410	-0.0067	0.0094

Table 4.12 Mixed model output of variables on average methylation at CGI amplicon 1 [BA: Brodmann Area; EC: Entorhinal Cortex; STG: Superior Temporal Gyrus; CSDD: Cornell Scale for Depression in Dementia; SLE: presence of a stressful life event; *: significant variable]

Variable		Coefficient	Std. Err.	Z	*p*	95% CI	
*CSDD score		0.0008	0.0003	2.4400	0.0150	0.0002	0.0015
*SLE		0.0103	0.0044	2.3200	0.0200	0.0016	0.0190
5HTTLPR S/S genotype		-0.0038	0.0057	-0.6700	0.5000	-0.0149	0.0073
*Age at death		0.0009	0.0003	3.0000	0.0030	0.0003	0.0014
Sex		0.0004	0.0045	0.0900	0.9310	-0.0085	0.0093
ApoE genotype		0.0004	0.0003	1.4900	0.1370	-0.0001	0.0010
Antidepressants		-0.0045	0.0053	-0.8500	0.3950	-0.0150	0.0059
Disease duration		0.0000	0.0005	0.0900	0.9300	-0.0010	0.0011
Brain region:	BA9	0.0084	0.0049	1.7100	0.0870	-0.0012	0.0181
	BA10	0.0041	0.0051	0.8100	0.4170	-0.0059	0.0141
	BA46	0.0038	0.0054	0.7100	0.4750	-0.0067	0.0144
	BA8	0.0105	0.0054	1.9500	0.0510	-0.0001	0.0211
	EC	0.0059	0.0058	1.0300	0.3040	-0.0054	0.0173
	*STG	0.0144	0.0061	2.3500	0.0190	0.0024	0.0264

Table 4.13 Mixed model output of variables on average methylation at CGI amplicon 2 [BA: Brodmann Area; EC: Entorhinal Cortex; STG: Superior Temporal Gyrus; CSDD: Cornell Scale for Depression in Dementia; SLE: presence of a stressful life event; *: significant variable]

4.6.2.3 Modelling the effects of predictor variables on SERT mRNA expression levels

Initially three mixed models of SERT mRNA levels were generated to test the association of the presence of a SLE, CSDD score and the presence of the 5HTTLPR S/S genotype separately. While the presence of a SLE was not associated with mRNA levels, the presence of the 5HTTLPR S/S genotype and CSDD scores were independently associated with SERT mRNA levels (β = -0.43, SE = 0.17, p = 0.009 and β = 0.027, SE = 0.010, p = 0.007, respectively). Additionally in all three models the mRNA levels of the BA9, EC and STG brain areas were significantly different from the baseline VC region (p = <0.001, for each area in each model). These differences in mRNA levels compared to the VC are clearly seen in a scatterplot of CSDD scores against mRNA levels (Figure 4.15).

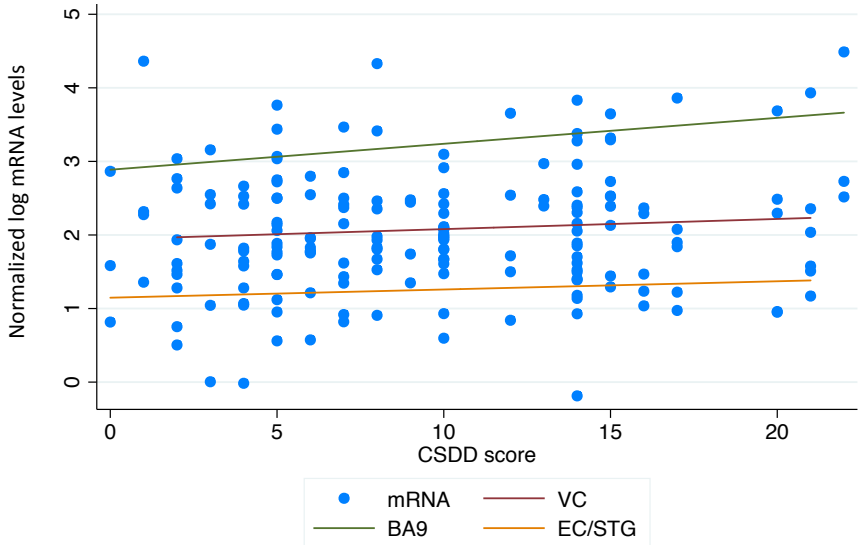


Figure 4.15 Scatterplot of CSDD scores against SERT mRNA levels and fitted lines for brain area [BA9: Brodmann Area 9; EC: Entorhinal Cortex; STG: Superior Temporal Gyrus; VC: Visual Cortex]

In a mixed model of SERT mRNA levels across the seven brain regions, controlling for the possible confounders of age, sex, ApoE genotype, antidepressant treatment and disease duration while testing the association of the presence of the 5HTTLPR S/S genotype, average CGI methylation and CSDD scores the following results were obtained as shown in Table 4.23. In this combined model the presence of the 5HTTLPR S/S genotype (β = -0.446, SE = 0.179, p = 0.013), average CGI methylation (β = 7.000, SE = 3.391, p = 0.039) and CCDS score (β = 0.034, SE = 0.011, p = 0.002) were all significantly associated with SERT mRNA levels. In addition the mRNA levels are significantly different from the VC at brain areas BA9, BA8, EC and STG (Table 4.14), with all having lower levels than the VC with the exception of the BA9 region. Also testing for an interaction between 5HTTLPR S/S genotype and methylation on mRNA levels, proved negative.

Variable		Coefficient	Std. Err.	Z	*p*	95% CI	
*Average CGI methylation		7.000	3.391	2.060	0.039	0.354	13.646
*CSDD score		0.034	0.011	3.040	0.002	0.012	0.056
*5HTTLPR S/S genotype		-0.446	0.179	-2.490	0.013	-0.796	-0.095
*Age at death		0.018	0.009	2.040	0.042	0.001	0.036
Sex		0.163	0.138	1.190	0.236	-0.107	0.433
ApoE genotype		0.012	0.009	1.350	0.176	-0.006	0.030
*Antidepressants		-0.425	0.162	-2.620	0.009	-0.743	-0.108
Disease duration		0.010	0.015	0.680	0.498	-0.020	0.041
Brain region:	*BA9	0.942	0.142	6.620	<0.001	0.663	1.221
	BA10	-0.268	0.153	-1.760	0.079	-0.568	0.031
	BA46	-0.186	0.166	-1.120	0.264	-0.511	0.140
	*BA8	-0.422	0.165	-2.550	0.011	-0.746	-0.098
	*EC	-1.284	0.194	-6.630	<0.001	-1.664	-0.904
	*STG	-0.871	0.214	-4.070	<0.001	-1.289	-0.452

Table 4.14 Mixed model output 4 of variables on SERT mRNA expression levels
[BA: Brodmann Area; EC: Entorhinal Cortex; STG: Superior Temporal Gyrus; CSDD: Cornell Scale for Depression in Dementia; SLE: presence of a stressful life event; *: significant variable]

Similarly modelling with the average CGI methylation at amplicon 1 and amplicon 2 as separate variables in a the one model, the same associations and differences in the brain regions are noted however the association of average CGI methylation at amplicon 1 with mRNA levels is significant (β = 7.45, SE = 3.3, p = 0.024) while the association at average CGI methylation at amplicon 2 is not significant (β = 2.19, SE = 2.07, p = 0.291) (Table 4.15). Again testing for an interaction between 5HTTLPR S/S genotype and CGI methylation at amplicon 1 on mRNA levels, proved negative.

Variable		Coefficient	Std. Err.	Z	*p*	95% CI	
*Average CGI methylation amplicon 1		7.446	3.297	2.260	0.024	0.984	13.907
Average CGI methylation amplicon 2		2.185	2.068	1.060	0.291	-1.867	6.237
*CSDD score		0.034	0.011	2.990	0.003	0.012	0.056
*5HTTLPR S/S genotype		-0.470	0.181	-2.600	0.009	-0.824	-0.115
*Age at death		0.018	0.009	2.020	0.044	0.001	0.036
Sex		0.155	0.139	1.110	0.265	-0.117	0.427
ApoE genotype		0.012	0.009	1.350	0.178	-0.006	0.031
*Antidepressants		-0.407	0.163	-2.490	0.013	-0.727	-0.086
Disease duration		0.012	0.015	0.780	0.438	-0.018	0.042
Brain region:	*BA9	0.968	0.142	6.810	<0.001	0.689	1.247
	BA10	-0.281	0.152	-1.850	0.064	-0.579	0.017
	BA46	-0.190	0.165	-1.150	0.251	-0.513	0.134
	*BA8	-0.430	0.165	-2.610	0.009	-0.753	-0.107
	*EC	-1.338	0.191	-7.000	<0.001	-1.713	-0.964
	*STG	-0.820	0.213	-3.850	<0.001	-1.237	-0.403

Table 4.15 Mixed model output 5 of variables on SERT mRNA expression levels
[BA: Brodmann Area; EC: Entorhinal Cortex; STG: Superior Temporal Gyrus; CSDD: Cornell Scale for Depression in Dementia; SLE: presence of a stressful life event; *: significant variable]

While treating the methylation levels at amplicon 1 and amplicon 2 separately in individual models neither amplicon 1 or amplicon 2 methylation levels on their own show any significant association with SERT mRNA expression levels, indicating that the better model accounting for all the data together is one where the data from amplicon 1 and amplicon 2 are treated as separate variables.

To summarise all the models on the different outcome measures tested, encompassing the Alzheimer's disease subset with clinical features presented in section 4.7, the significant associations are illustrated in Figure 4.16.

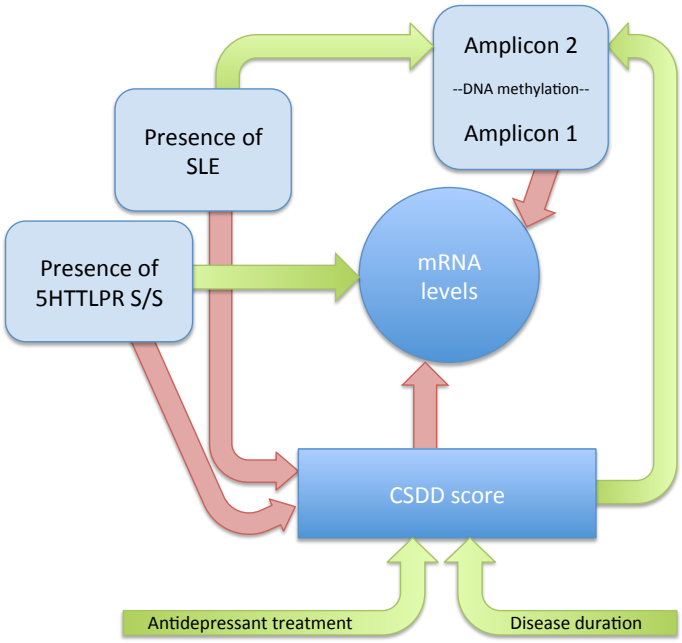


Figure 4.16 Illustration of main significant associations with the different outcome measures tested [green arrows: significant association in the expected direction; pink arrows: significant association in the unexpected direction]

4.7 Discussion

The observed methylation levels at the SERT CGI are extremely low and below the technical detection level of the Sequenom MassARRAY EpiTYPER platform, which is around 5% (Coolen et al. 2007). Given that all the data are below 5% and that any measurable differences are very small and not likely to be significant, the analysis in this chapter was purely exploratory. Where associations are noted as being significant these are not robust enough to withstand corrections for multiple testing and are thus only presented as tentative findings. Further still considering the effect sizes and differences in DNA methylation associated with the variables tested in this

chapter are well below 0.5% there is likely to be very little biological plausibility to results presented.

Measurement of average methylation levels at the SERT CGI across seven different brain areas from post-mortem tissue of AD and controls, shows variable levels at the different brain areas, with the BA8, BA9 and BA46 areas having significantly higher levels than the VC control brain area (Table 4.1). More specifically average methylation levels of CGI amplicon 1 do not differ significantly from the VC (Table 4.2), while average methylation levels of CGI amplicon 2 at brain areas BA9, BA46 and STG are significantly higher than the VC (Table 4.3).

Average methylation levels of CGI amplicon 2 reveal a significant relationship with age, whereby each year in age denotes a 0.0006% increase in methylation. Though this is based on only data from individuals at death and not from the same individual over different time points, it gives an indication as to a change in SERT methylation over time. In addition females displayed on average 0.008% more methylation at CGI amplicon 2 than males at the frontal cortex areas, this is in line with previously described sex differences, measured in lymphoblast cell lines (Philibert et al. 2008).

While AD status is significantly associated with the average methylation levels of the SERT CGI in brain area BA10, with AD patients having on average 0.009% higher levels compared to controls, CpG fragment b22 (within amplicon 2) presented a similar significant relationship with AD status. Though when this relationship was tested with SERT mRNA expression data there was no significant association of either AD status or Braak staging with mRNA levels. Whereas average CGI amplicon 1 methylation levels are positively associated with mRNA levels.

However the observed relationship of methylation with mRNA levels is in the opposite direction to that expected and as described previously in both buccal and lymphoblast cell lines along with *in vitro* experiments of the SERT promoter (Philibert et al. 2008; Olsson et al 2010; Wang et al. 2012). Although, the authors of these studies do acknowledge the various limitations in using these peripheral tissues and that they may not directly be extrapolated to the human brain, they all demonstrate lower SERT CpG methylation to associate with higher mRNA expression. From these studies the only corresponding CpG site that could be accurately mapped and tested in the current dataset was the CpG unit at fragment b34 reported by Philibert et al. (2008), though when tested within this dataset the association of CpG b34 methylation with mRNA expression was in the expected negative direction, it did not reach significance.

Further still, Philibert et al. (2008) highlight a potential issue with the use of different mRNA probe sets in accurately measuring the levels of SERT mRNA expression. Whereas the mRNA data used in this chapter was derived from a probe set covering the SERT exon 8 and 9 boundary, Philibert et al. (2008) used probe set covering the exon 1 and 2 boundary, suggesting that a probe set covering the SERT exon 8 and 9 in addition to measuring full-length transcripts, will also recognize partial transcripts that result from promiscuous transcription along the gene, which are not thought to be regulated by the promoter or be translated into a functional protein and may thus be biologically irrelevant (Philibert et al. 2008).

Consistent with the observations of lower levels of mRNA expression in S allele carriers (Lesch et al. 1996), the data presented noted those AD patients with the S/S genotype as having on average less mRNA expression levels than those with either of L/L, L/S genotypes at the frontal cortex areas. Although the data are unable to establish if there was a interaction between 5HTTLPR S/S genotype and the presence of a SLE in relation to depression or mRNA levels in AD.

Interestingly, while there was no effect of 5HTTLPR genotype on any measure of CGI DNA methylation, both presence of SLE and higher depression scores are significantly associated with higher levels of CGI amplicon 2 methylation. Further still though there was no observation of higher levels of methylation in S/S genotype individuals compared to L/L individuals, such as that noted in recent studies looking a SERT in blood (Kinnally et al. 2010; Alasaari et al. 2012), these studies go on to show the effect of stress and SERT DNA methylation on their outcome behavioural measure maybe independent of 5HTTLPR genotype (Kinnally et al. 2010; Alasaari et al. 2012). Similarly the present data could be interpreted as indicating a separate role for SLE in modifying risk to depression in AD through DNA methylation aside from solely involving a interaction with the 5HTTLPR polymorphism as cited in the literature (Caspi et al. 2003), nevertheless the methylation data do not establish a clear relationship with mRNA levels, as noted above.

However a limitation of the current data is that the measures of SLE used are ad hoc because we did not have proper lifetime assessments based on a validated inventory, such as that described by Brugha et al. (1985), the SLE measure used here does not capture all possible SLEs and thus likely limited in truly reflecting this environmental stressor. Furthermore it has been suggested that early onsets of depression may be more influenced by gene x environment interactions, such as that described by Caspi et al. (2003) than late-onset episodes (Gillespie et

al. 2005) and so this paradigm may not be appropriate to the study of late-onset depression in AD.

DNA methylation and mRNA expression levels can differ within the brain and such differences along with SERT density are likely to extend to other parts of the CNS and peripheral tissues, possibly accounting for some of the discrepancies between studies (Gardner et al. 2009). In addition some of these differences may arise through an interaction between adverse early life experience and stress during adulthood leading to increased SERT mRNA expression within a subpopulation of serotonergic neurons (Gardner et al. 2009; Arango et al. 2001). Furthermore these conclusions maybe complicated by the existence of subgroups within the dichotomous 5HTTLPR S/L allele paradigm and additional SNPs within its sequence (Avula et al. 2011; Parsey et al. 2006), or a variable number of tandem repeat (VNTR) polymorphism STin2, in intron 2 with regulatory potential, which has also been associated with aggression in AD (Ueki et al. 2007).

Potentially, any of these issues may also account for the unexpected positive direction of the association of depression with mRNA levels. Though significantly AD patients with the SERT 5HTTLPR L/L, L/S genotypes have higher depression scores compared to those with the S/S genotype, which is in line with a much wider dataset of clinical/behavioural data within which this study is nested. This association of genotype alone with depression is akin to findings demonstrating the L allele to be associated with increased risk for anxiety or aggression and psychosis in AD (Tsang et al. 2003; Quaranta et al. 2009; Sukonick et al, 2001; Sweet et al. 2001). Also these preliminary findings demonstrate an association of L/L, L/S genotypes with depression whereas; others have previously found no effect of 5HTTLPR genotype on depression in AD (Li et al. 1997; Assal et al. 2004; Micheli et al. 2006; Pritchard et al. 2007; Ueki et al. 2007; Grünblatt et al., 2009).

However in contrast to the current work all of these studies looking at genotype and depression in AD are based on much larger sample sizes, highlighting one of the main limitations of the present findings in that they are likely to be underpowered and sampling bias can not be ruled out. As exemplified by the inability to detect any consistent association of SLE and CSDD scores, mentioned in section 4.6.2. Other inherent limitations of such post-mortem studies include those already described in the discussion from chapter 3 (section 3.7), with regards to measurements of DNA methylation in post-mortem brain homogenates; and the difference in time between baseline interviews, when clinical/behavioural data were acquired to when the

measurement of DNA methylation and mRNA was made from post-mortem tissue, in that these measurements might not truly reflect the behavioural phenotype at death.

Future studies would benefit from behavioural assessments of patients as far up to death as practically feasible and based on the recent observation that peripheral SERT methylation could be a marker of central serotonin function (Wang et al. 2012) they may even incorporate mid to late life SERT DNA methylation and mRNA quantification in blood collected over several time points to better asses the temporal relationship between these measures, SLEs and depression in AD.

APP GENE REGION DNA METHYLATION ANALYSIS BY BISULPHITE SEQUNCING

5.1 Introduction

AD pathology

The observation of extracellular plaques in the brain during postmortem examination is one of the characteristic pathological features of Alzheimer's disease. Plaques are primarily composed of aggregations of amyloid beta peptides (Aβ40/42). The Aβ peptide is derived from a larger precursor, type-1 transmembrane protein called amyloid protein precursor (APP). Aβ is generated from APP through sequential cleavages first by β-secretase and then by γ-secretase complex along with a slightly smaller soluble fragment called sAPPβ. Alternatively, APP can be cleaved by α-secretase within the Aβ domain to release soluble sAPPα and preclude Aβ generation. Differential processing of APP in the normal verses disease state is thought to play a pivotal role in AD, whereby an imbalance of Aβ production and Aβ clearance is seen as the initiating event in the pathogenetic cascade eventually leading to neurodegeneration and cognitive decline associated with AD (Hardy & Selkoe 2002).

This is supported by the fact that most of the genetic mutations causing early-onset familial AD (EOFAD) are missense mutations, located in or flanking the Aβ encoding sequence of the APP gene or within the presenilin (PSEN1 and PSEN2) genes, which encode the proteolytic subunit of the γ-secretase complex (Bertram et al. 2007, www.alzgene.org). While EOFAD accounts for only a small proportion of AD (<5%), and the majority of the AD cases are late-onset sporadic AD (LOAD), it is reasonable to suggest that LOAD may represent a variation of this pathogenetic cascade involving APP, leading to Aβ accumulation later in life. Since the factors involved in APP processing are crucial for Aβ accumulation, it is conceivable that the amount of APP substrate made available by transcriptional regulation of APP plays an important role in AD pathology.

Down's syndrome & APP duplication

A number of different observations indicate that altered expression of the APP gene may play an important role in Aβ deposition and AD. Firstly in Down's syndrome (trisomy 21), over-expression of APP due to triplication of chromosome 21 on which the APP gene resides, leads to extensive AD neuropathologic changes including Aβ amyloid plaques (Wisiewski et al. 1985).

Most individuals with Down's syndrome develop clinical and pathological Alzheimer's disease after the age of 40 years (Oyama et al. 1994; Coppus et al. 2006; Bird 2010). While the overexpression of APP has not been observed in the developing brain up to 18 weeks gestation, postmortem adult Down's syndrome brain shows a marked increase in APP relative to controls (Cheon et al. 2008). Others have suggested a 4-fold overexpression of APP in Down's syndrome, results in a 50 year decrease in onset age of AD in these individuals compared with the normal population (Rumble et al. 1989). Although the possible contribution of other genes located on chromosome 21 to the process of Aβ accumulation in Down's syndrome individuals cannot formally be excluded, there are reports of smaller duplications involving the APP region that do not cause Down's syndrome but that contain an additional copy of APP are sufficient to cause AD (Sleegers et al. 2006).

Other duplications involving the APP region and leading to overproduction of APP have been identified in European families with autosomal dominant early-onset Alzheimer's dementia with cerebral amyloid angiopathy (Rovelet-Lecrux et al. 2006, 2007; Cabrejo et al. 2006; Sleegers et al. 2006). Two unrelated Japanese EOFAD families with duplications encompassing the APP region were also reported with accompanying increases in APP mRNA in blood compared to age matched controls (Kasuga et al. 2009). Some of the smaller duplications across the APP gene are reported to extend maximally over ~0.7Mb, where one duplication solely covers the APP gene (Sleegers et al. 2006) and another, a region ranging from APP to the neighbouring CYYR1 gene (Kasuga et al. 2009). Furthermore Hooli et al. (2012) have reported a family with the smallest duplication to date covering just the APP gene (0.38Mb). However they also note incomplete penetrance of this mutation as it is carried by only two of the three affected siblings in this family (onset ages 49 and 52 years), whereas a third affected individual (onset age 70 years) showed a diploid copy status in this region and while two further unaffected siblings (ages 69 and 74 years) also showed no evidence for the duplication (Hooli et al. 2012). Other protective factors or single nucleotide polymorphisms (SNPs) in the sequence flanking the duplication could possibly be involved.

APP promoter polymorphisms

Variants both within the proximal promoter and further into the 5' region of the APP gene have been examined for association with AD risk that may lead to over expression of APP. Initial investigations concluded that three APP promoter polymorphisms (-209C/G, -9G/C and +37G/C, +1 transcription start site) are unlikely to contribute strongly to AD susceptibility or to cause major differences in APP expression, though the +37C allele (rs459543) was over represented

among patients with AD compared with elderly controls (OR, 1.57; 95% CI, 1.08- 2.27) this was not significant after adjusting for age, sex, and education (OR, 1.41; 95% CI, 0.93-2.12) (Liddell et al. 1995; Athan et al. 2002; Lv et al. 2008).

On the other hand Theuns et al. (2006), have shown that rare genetic variants in the 5' APP regulatory region that increased APP transcriptional activity in neuronal cells are associated with AD (Theuns et al. 2006). Using two independently ascertained cohorts from the Netherlands and Belgium. They identified six genetic variants in AD patients only, of which five (-118C/A, -369C/G, -534G/A, -479C/T and -371G/A) predicted a change in transcription factor binding site (TFBS) affinities. Using a luciferase reporter gene assay, they observed for three (-118C/A, -369C/G and -534G/A) of these five variants a significant neuron-specific increase of APP transcription in human neuroblastoma SH-SY5Y cells of nearly 2-fold. While one other variant (-479C/T) showed only a 1.2-fold increase, for which the contribution to AD risk was reported to be more pronounced at later age due to modulating factors (Brouwers et al. 2006). In addition four of these variants are located in a 200bp fragment (-540 to -340) that is highly conserved between species (>95%), which is indicative of a functional role for this sequence (Brouwers et al. 2006). Together these two studies provided evidence that AD mutations in APP regulatory sequences do occur with onset ages inversely correlated with levels of APP expression (Theuns et al. 2006; Brouwers et al. 2006).

However the association of these three functional variants (-369C/G, -479C/T, and -534G/A) (Theuns et al. 2006; Brouwers et al. 2006) along with another two variants (-1023T/C and -3830C/T) that have been reported to have an effect on APP gene expression (Lahiri et al. 2005) with the risk of developing AD have proved negative in a French cohort (Guyant-Marechal et al. 2008). Though the same study reported an association between rs463946 (-3102 G/C) and AD that was confirmed in a replication sample of a similar size (Guyant-Marechal et al. 2008), a much larger study in a Caucasian cohort has failed to confirm the positive associations of rs463946
(-3102 G/C), rs459543 (+37G/C), -369C/G, -479C/T, and -534G/A sequence variants with AD risk (Hooli et al. 2012). Another study in the Chinese Han population identified two additional polymorphisms, rs466433 (-877T/C) and rs364048 (-955A/G) to associate with AD (Lv et al. 2008). The -877T and -955A alleles were observed to be over-represented in LOAD patients when compared to controls. Furthermore, the authors noted the variants -877T/C and -955A/G to be in strong linkage disequilibrium and constitute a relatively risky -877T/-955A and a relatively protective -877C/-955G haplotype, with luciferase reporter assays showing the -

877T/-955A construct to have a 4-fold increase in transcriptional activity than -877C/-955G (Lv et al. 2008).

Although none of the currently published GWASs in AD (Bertram et al. 2007, www.alzgene.org) have so far reported significant association between risk for AD and common variants in or near the APP gene and though this may rule out common sequence variation as a contributing factor to risk for LOAD, the observations of increased APP expression through genomic duplication or rare mutations in the APP regulatory region that lead to EOFAD lends support to the hypothesis that APP expression and Aβ are central to AD aetiology. Furthermore expression microarray studies show increased expression of APP in hippocampus (Liang et al 2008) and entorhinal cortex (Guttula et al. 2012) areas of AD brain relative to controls and while these experiments indicate an overall upregulation of APP, it is worth noting that different APP isoforms exist for which mRNAs are altered in AD brain.

Major Isoforms of APP in brain

As already mentioned the APP gene is on chromosome 21 and contains 18 exons spanning over 0.3Mb. While several isoforms of APP have been observed in humans, ranging in length from 305 to 770 amino acids; alternative splicing of exons 7 (which encode a 56-amino acid Kunitz-type proteinase inhibitor (KPI) domain) and 8 (a 19- amino acid domain that shares sequence identity with the OX-2 antigen) results in three major isoforms. The longest isoform, APP770, contains both the KPI and the OX-2 domains, whereas APP751 contains only the KPI domain. The shortest isoform, APP695, lacks both domains. In the brain, APP695 is expressed at high levels, and the APP751/770 isoforms are expressed at significantly lower levels (Koo et al. 1990),

In situ hybridization studies have demonstrated higher levels of APP mRNA in brains of patients with AD (Higgins et al. 1988; Cohen et al. 1988; Lewis et al. 1988). While Higgins et al. (1988) used probes for total APP mRNA and observed an increased signal in AD hippocampus, others have used probes for total APP mRNA and APP695 and found that APP695 mRNA increases selectively in surviving nucleus basalis and locus ceruleus neurons but not in neurons of the hippocampal subiculum, occipital cortex, or basis pontis (Cohen et al. 1988). More specifically there is cell specific transcriptional and post-transcriptional regulation of APP gene expression in CNS cell types, where total levels of APP mRNA and APP695 mRNA are higher in neurons while glial cells contain higher levels of KPI-containing mRNAs (Menéndez-González et al. 2005). Although there are regional differences in the brain it has been suggested that the balance

between the KPI- and non-KPI containing isoforms may be an important factor in AD (Moir et al. 1998; Golde et al. 1990).

While measuring a number of different cell specific markers and core elements of the APP metabolome a shift in AD temporal neocortex away from the neuronal APP695 isoform towards the KPI-containing isoforms was noted to occur specifically within the context of gliosis (Matsui et al. 2007). Though the association between total APP protein and mRNA levels was not significant, APP-KPI protein levels significantly correlated with APP-KPI mRNA levels along with a elevation in soluble Aβ peptide levels that was suggested to lead to accelerated amyloidogenesis in sporadic AD based on Aβ production by glia cells (Matsui et al. 2007). Although upregulation of KPI-containing isoforms of APP resulting from excitotoxic insult, are thought to exert important neuroprotective functions against further neuronal loss (Mattson 1997), these changes in the ratios of APP isoforms with and without KPI domains may result in alterations of APP processing that favour the formation of amyloid.

APP gene regulation

The APP gene is subject to complex regulation, including both distal (Lahiri & Robakis 1991; Lahiri et al. 1999) and proximal (Ge et al. 2004) regulatory elements. Sequence analysis and functional promoter assays of Rhesus monkey and human APP promoter have been used to identify several elements that are likely to be important in controlling expression (Lahiri 2004). For instance sequence extending from about -600 to -460 bp (+1 being the transcription start site, TSS) acts as a positive regulator, while a second segment extending from -450 to -150 bp acts as a negative regulator (Lahiri & Robakis 1991). Additionally upstream regulatory elements that control the activity of the human APP promoter are from -2257 to -2234 bp and from -489 to -452 bp (Lahiri 2004). The former of these two sequences is a negative regulator of transcription, referred to as the upstream regulatory element (URE), binding an unknown transcription factor (TF), expressed in a restricted number of cell types and different regions of the human brain (Lahiri et al. 1999). While the proximal promoter region is devoid of a functional TATA box, transcription initiation is regulated by a strong initiator element (LNR) surrounding the TSS (-7 to +7 bp) (Quitschke et al. 1996; Lahiri et al. 2004). The sequence of both the LNR and a upstream element (UE, -30 to -12 bp) are associated with DNase I-protected domains, suggesting sequence-specific binding of nuclear factors (Quitschke et al. 1996).

With multiple transcription factors predicted to interact with the APP promoter region, two main GC-rich elements are thought to drive APP promoter activation. The sequence within -93

to -82 bp, termed APB-β domain is thought to contribute most to APP promoter activity in both neuronal and non-neuronal cells and is shown to bind the CCCTC-binding factor (CTCF), transcription regulator with diverse roles in regulation of cell growth, differentiation and apoptosis, enhancer-blocking activity and control of imprinted genes (Theuns & Broeckhoven 2000; Burton et al. 2002; Ohlsson et al. 2001). The remaining APP promoter activity is accounted for mainly by the APB-α domain (-65 to -41 bp) and is mediated by binding of stimulating protein 1 (SP1) and the upstream stimulatory factor (USF at -48 to -41 bp). SP1 belongs to a family of zinc finger protein transcription factors that includes other members such as SP2, SP3 and SP4 (Kadonaga et al. 1987; Hagen et al. 1995). SP1 is known to be an important accelerator of APP mRNA production (Pollwein et al. 1993) and that the overall distribution and cellular localization of SP1, APP, and Aβ are similar and neuronal in origin (Brock et al. 2008). In addition SP1 activation of the APP gene may be subject to perturbation by environmental metal exposures (see below, Wu et al. 2008; Zawia et al. 2009). An interaction model for APP transcriptional activity based on DNA looping has also been suggested, where a novel SP1-like protein forms a homodimer tethering the USF and 5'-flanking SP1 sites, and a upstream AP1 site with flanking GC-rich/box motif (Querfurth et al. 1999). A heat shock element (HSE) at position -317 is also noted for transcriptional activation of APP through the binding of heat shock factor-1, possibly leading to a 1.6-fold increase in transcription (Dewji & Do 1996).

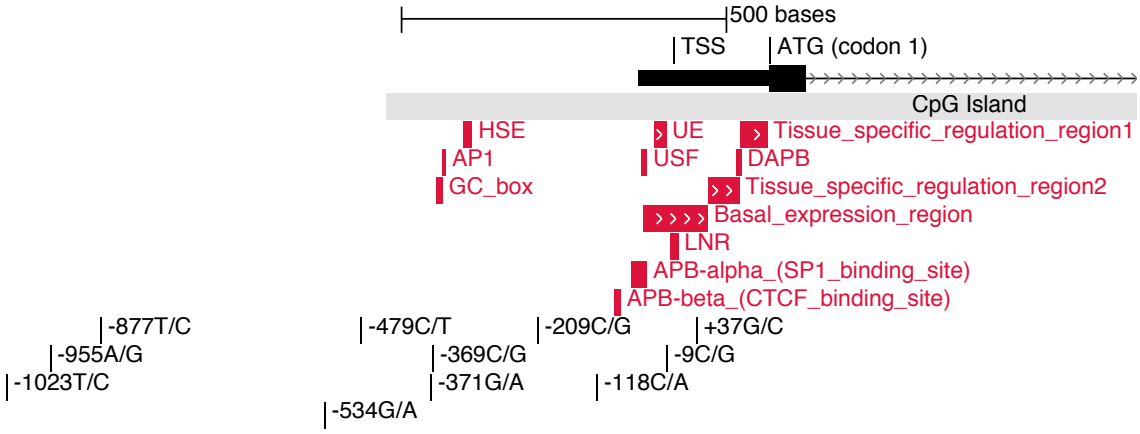


Figure 5.1 Schematic of APP gene promoter and regulatory features
[APP gene exon 1 with transcription start site (TSS) and first translation codon (ATG) indicated; grey box: CpG island (CGI); red boxes: characterised regulatory sequences; black vertical bars: promoter polymorphisms]

The structure and expression pattern of the APP gene are highly conserved between species (Yamada et al. 1989; Maloney et al. 2004; Lahiri 2004), with ~80% homology of the promoter regions of rodents and primates to that of the human APP (Wirak et al. 1991; Podlisny et al. 1991). They all lack typical TATA box, have a high GC content, and share TFBSs mentioned above, suggestive of a major role for these TFs in APP regulation in vivo (Theuns & Broeckhoven

2000). Though, deletion mapping of the APP promoter demonstrated that <100 bp upstream of the TSS are sufficient for high levels of expression in numerous cell types (Lahiri & Robakis 1991; Quitschke & Goldgaber 1992; Pollwein et al. 1993), more recent studies of the Rhesus monkey APP promoter in transfected cell lines, showed that a 30 bp (-76 to -47 bp) region acts as a negative regulatory element in NB (human SK-N-SH neuroblastoma cells) and a positive regulatory element in PC12 (rat pheochromocytoma) cells. While a upstream region (> 1 kb from the TSS) is necessary for activity in HeLa (human cervix epithelial) and C6 (rat glial) cells, but not for NB cells (Ge et al. 2004). Further reporter gene studies have shown that the human APP gene promoter basal expression element is at -46 to +54 bp and two cell-type-specific elements at +54 to +100 bp and +100 to +144 bp suggesting the gene's 5' (un-translated region) UTR and proximal promoter region interact as a unit to regulate gene expression at both the transcriptional and post-transcriptional levels (Lahiri et al. 2005). More recently another nuclear factor-binding domain designated as DAPB with a recognition sequence located between position +96 and +105, also within the 5' UTR was shown to exert its effect through a post-transcriptional mechanism (Vostrov et al. 2010), while a micro-RNA sequence in the 3'UTR was noted to down regulates APP protein levels in human cell culture (Long & Lahiri 2011). Though there are both post-transcriptional and translational mechanisms at play in the regulation of APP expression, DNA methylation of the APP gene promoter may also possibly affect APP transcription.

APP & DNA methylation

Global changes of DNA methylation in animal models have been shown to affect specific Aβ related genes, leading to changes in expression of these genes and accumulation of Aβ. Exposure of new born monkeys to metallic lead (Pb) upregulates the expression of APP, BACE1, as well as their transcriptional regulator Sp1 in various cortical areas at old age accompanied by a decrease in DNA methyltransferase activity, and a higher level of oxidative damage to DNA, indicating that epigenetic changes in early life can interrupt the expression of AD related genes and promote DNA damage and pathogenesis (Wu et al. 2008; Zawia et al. 2009).

In addition DNA methylation is achieved by adding a methyl group from S-adenosyl methionine (SAM) and is demethylated into S-adenosylhomocysteine (SAH; an inhibitor of methyltransferase), which hydrolyzes to produce homocysteine (HCY). The reversible nature of this reaction means increased HCY would lead to SAH accumulation and cause DNA hypomethylation. A deficiency of folate and vitamin B12 can cause a reduction in SAM level (Fuso et al. 2011; Fuso & Scarpa 2011). For example, it has been shown that increased plasma

HCY and decreased SAM levels are found in AD patients, which in turn could lead to APP promoter hypomethylation and upregulation (Schulz 2007; Haan et al. 2007; Coppedè 2010). Furthermore, mice reared on a diet deficient in choline, essential for the biosynthesis of SAM, exhibit altered global and gene specific DNA methylation patterns and APP is one of the genes showing over expression with choline deficiency (Niculescu et al. 2006).

Induced oxidative stress models of human neuroblastoma SH-SY5Y cells have demonstrated hypomethylation of the APP promoter and increased Aβ production (Guo et al. 2011; Gu et al. 2013) and though such induced differences in cell models are also not specific to APP, it was demonstrated that a 27% reduction of APP promoter methylation can result in a 3-fold increase of APP mRNA expression (Guo et al. 2011). APP is also up regulated by hypomethylation of its promoter as seen in SAH (S-adenosylhomocysteine) administered mouse microglial cells (Lin et al. 2009) and in rat neural cell lines grown in the presence of the demethylating agent 5-azacytidine (Ledoux et al. 1994).

As previously mentioned in Chapters 1 and 3, the interconnected processes of aging, oxidative stress and a diet related decreased SAM:SAH ratio have all been implicated as risk factors for AD, accompanied by a decrease in DNA methyltransferase activity and result in global hypomethylation of DNA (Fuso & Scarpa 2011). Given these observations of hypomethylation at the APP promoter within the context of aging it is conceivable that increased expression of the APP gene with aging in LOAD may to some extent be regulated by its methylation status.

This hypothesis is partially supported by bisulphite sequencing experiments that identified at least 13 potential methylation sites in the region -236 to -101 bp of the human APP promoter where ~26% of these cytosines were more frequently methylated in healthy individuals between the ages of 35 and 70 years compared to ~8% in those aged 74-90 years (Tohgi et al. 1999a ; Nagane et al. 2000). This age-related reduction in methylation, which would be consistent with an increase in APP expression with aging, was more prominent in GC-rich elements matching SP1 DNA-binding sites (-207 to -182 bp).

The relationship between APP methylation at sequences -3500 to -450 bp (from the TSS) and APP gene expression has also been explored (Rogaev et al. 1994). Using a methylation restriction approach and several different human tissues, including multiple brain areas; Rogaev et al. (1994) reported a CpG site around -1184 bp with variable methylation that crudely correlated with differences in APP transcription among different tissues, with the neocortex showing higher levels of methylation compared to the cerebellum and other tissues. Though

these studies (Rogaev et al. 1994; Tohgi et al. 1999a ; Nagane et al. 2000) did not include AD patients, another study using the methylation restriction approach has reported a single CpG to be hypomethylated in AD temporal cortex from a single patient compared to a control and a individual with Pick's disease (West et al. 1995). However the authors of this last study do not report the precise location of this differentially methylated CpG site.

However more recent studies on human brain tissue using more sensitive DNA methylation detection techniques such as matrix-assisited laser desorption ionization time-of-flight (MALDI-TOF) (Wang et al. 2008; Barrachina and Ferrer, 2009) or modified Sanger sequencing (Milici et al. 1990; Brohede et al. 2010) have failed to replicate the APP hypomethylation observed in the studies mentioned above (Rogaev et al. 1994; West et al. 1995; Tohgi et al. 1999a ; Nagane et al. 2000). Though two of these studies with negative findings were carried out on both AD cases and controls (Wang et al. 2008; Barrachina and Ferrer, 2009) the other two were either centered on samples from AD patients (Brohede et al. 2010) or solely from healthy individuals (Milici et al. 1990).

Also though the DNA methylation detection techniques used in some of these studies are sensitive and widely used (Tohgi et al. 1999a ; Nagane et al. 2000; Wang et al. 2008; Barrachina and Ferrer, 2009; Brohede et al. 2010) while those used by others (Milici et al. 1990; Rogaev et al. 1994; West et al. 1995) are less sensitive, it may be difficult to accurately detect small changes in DNA methylation in vulnerable cells as they relate to chronic diseases in which small modifications may be sufficient for disease progression.

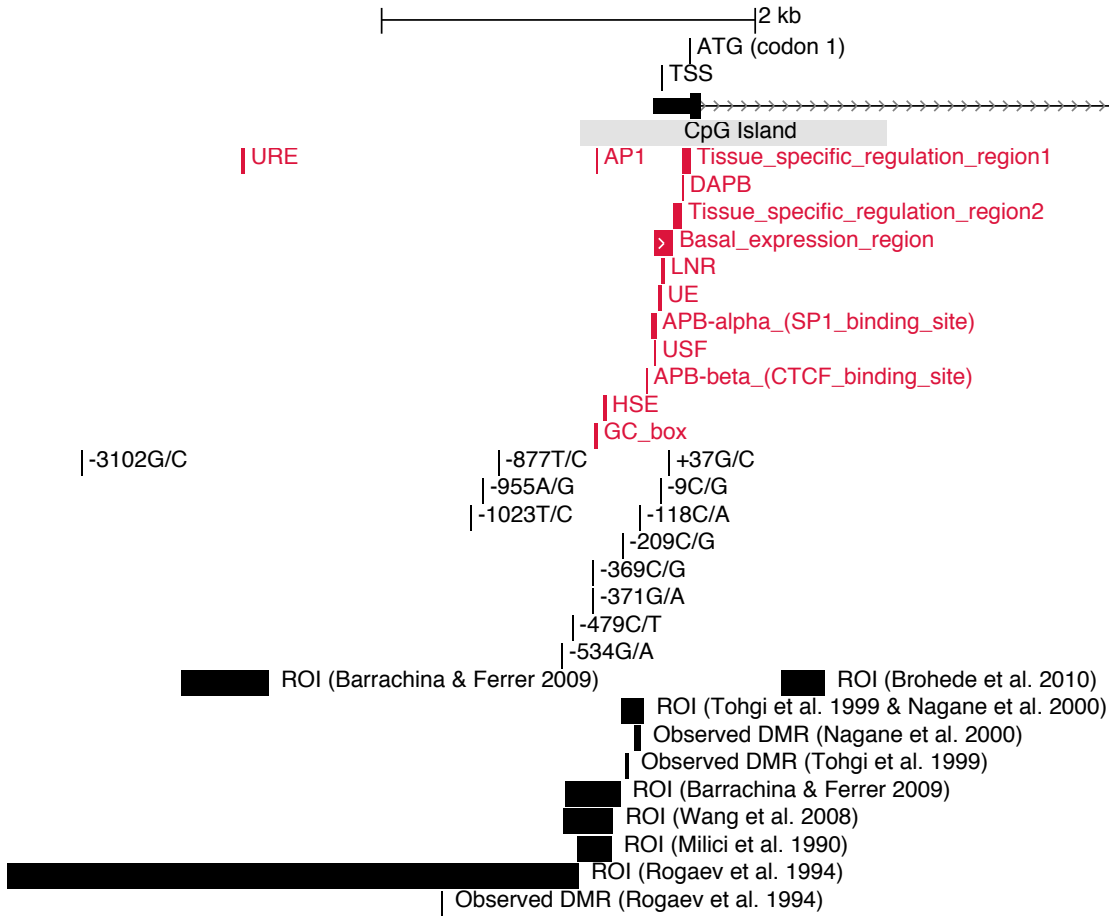


Figure 5.2 Schematic of APP gene promoter and promoter ROI sequences studied to date [APP gene exon 1 with transcription start site (TSS) and first translation codon (ATG) indicated; grey box: CpG island (CGI); red boxes: characterised regulatory sequences; black vertical bars: promoter polymorphisms; black boxs: ROI or DMRs studies to date]

Mixed results of these studies exploring methylation at the APP promoter region could also relate to differences in either the brain area chosen for investigation and/or the region of the APP promoter that was examined by individual studies. For instance the differentially methylation region around the 5' sequence -1184 bp from the TSS reported by Rogaev et al. (1994) is outside the CpG island (CGI), whereas the methylation pattern for a CpG site (-372) closest to the TSS was unmethylated in all tissues investigated. This latter result is more consistent with the other studies that only investigated short sequences in the proximal promoter region. However, given the complexity of APP transcriptional regulation as noted above and more recent reports showing that differentially methylated regions (DMRs) at nearby sequences (such as CGI shores ±2kb) can be important for the tissue-specific expression of genes (Irizarry et al. 2009); it is possible that there maybe further DMRs relating to APP expression that lie outside the region of interest (ROI) sequences investigated to date.

Gene body methylation

Not only has the influence of DNA methylation on transcription been demonstrated to extend into the first exon of a gene (Brenet et al. 2011), but also a growing number of reports suggest methylation within the gene body (intragenic methylation) could play a role in the transcriptional process. For instance Maunakea et al. (2010) have shown methylated intragenic regions to correlate with higher levels of gene transcription in human brain, which is also documented in cell transfection experiments (Bauer et al. 2010). Work by Maunakea et al. (2010) also suggests intragenic methylation may be a mechanism that regulates the use of alternative promoters. While it is linked to the modification of the transcriptional efficiency of RNA polymerase II and the transcriptional complex (Lorincz et al, 2004); intragenic DNA methylation was demonstrated to prevent CTCF binding, facilitate RNA polymerase II elongation and promote exon skipping in the protein tyrosine phosphatase, receptor type, C (PTPRC) gene (Shukla et al. 2011). In addition, Hodges et al. (2011) demonstrated hypomethylated regions (HMRs) that extended into the shore regions around promoters, as well as intragenic HMRs. Although the functions of these HMRs are unknown, the intragenic regions are associated with transcription factor binding sites, which suggest a biological role (Shenker & Flanagan 2012). More recently exploration of methylation differences in cancer further away from the CGI has shown enhancer methylation to be more closely associated with gene expression alterations than promoter methylation (Aran et al. 2013).

In an attempt to integrate genetic epigenetic and expression data Gibbs et al. (2010) explored the association of SNPs with, expression quantitative trait loci (eQTL) and CpG methylation quantitative trait loci (methQTL) in human brain and though they observed little co-association between methQTLs and eQTLs, possibly as a result of poor power to detect such effects (Gibbs et al. 2010), they found that only 18% of significant methQTLs were at CpG sites within CGIs whereas the majority were found to occur at CpG sites outside of the CGis. They also showed that cis-methQTLs are enriched at sequences 45bp from the CpG site (Gibbs et al. 2010). These findings along with studies looking at gene body methylation suggest that exploring methylation outside of the core CGI is warranted as it may also be involved in gene regulation.

Base-pair resolution CpG methylation analysis

Capillary electrophoresis based bisulphite sequencing has been considered the gold standard for methylation analysis because of its clear readout and single-base resolution (Frommer et al. 1992) but it can only be applied to relatively small regions in a small number of samples due to the labour intensive nature of the method. Whereas targeted parallel sequencing allows much greater throughput of samples and regions that can be interrogated simultaneously. Strategies for targeted ROI sequence capture rely upon PCR, in-solution capture, or microarrays to amplify or isolate regions of interest. However one disadvantage with the array-based methods is the inability to interrogate multiple closely apposed CpGs, while in-solution capture can show significant variability in the capture efficiency for each of the probes used. Although in-solution capture is one of the more widely adopted strategies, currently only a small number of samples can be multiplexed using proprietary methylated sequencing adaptors which are required for the parallel sequencing reaction. Bisulphite conversion coupled with PCR can overcomes these limitations, but single-plex PCR of multiple amplicons can be a time-consuming process (Fouse et al. 2010). To address this issue, recently two microfluidics technologies have become available which perform high throughput PCR amplification that can be easily applied to downstream parallel targeted sequencing, firstly the microdroplet PCR amplification system developed by RainDance Technologies (Tewhey et al. 2009) and secondly the Access Array 48.48 PCR system from Fluidigm and while the use of the RainDance platform with bisulphite sequencing has been established (Komori et al. 2011), it can be cost prohibitive owing to the cost of specialized primers and depending on the number of amplicons that are tested. Both technologies are claimed to significantly decrease PCR amplification bias (Tewhey et al. 2009; Hollants et al. 2012) and provide more evenly amplified products that can be directly applied to sequencing.

To summarize APP mutations and the neuropathology of individuals with three copies of this gene provide clear evidence that an alteration in either the amount of Aβ produced or the amino acid sequence of Aβ is sufficient to cause AD and strongly supports the notion that increased expression of APP can cause an increase in concentration of Aβ40/42 peptides and subsequent Aβ deposition that is critical for AD pathogenesis. While APP expression produces a heterogeneous product in the form of three main isoforms, APP gene regulation and transcription is complex involving both proximal and distal regulatory sequences that are likely to be influenced by DNA hypomethylation seen in the aging brain of LOAD patients. Though a number of studies presenting mixed results have already explored DNA methylation differences at a small select region of the APP promoter via varying methods, no study to date has explored

DNA methylation across the whole CGI and within the gene body. Emerging evidence suggests that intragenic methylation may influence transcription. Finally targeted base-pair mapping of individual CpG sites across a given gene region for a large number of samples is now feasible with the recent advances in microfluidics combined with parallel sequence.

To further investigate the potential role of APP gene DNA methylation in AD we aimed to carry out highly parallel bisulphite sequencing of the entire CGI, CGI shores and coding exons with intronic sequence boundaries. Along with a selected number of potentially regulated sequences, identified using data from the Encyclopaedia of DNA Elements (ENCODE) project. Considering that the APP promoter sequence is highly conserved (Maloney et al. 2004; Lahiri 2004), transcription factor binding sites that reside in these sequences and other conserved enhancer-associated elements may present regions worth investigating for DMRs. At the same time that highly parallel bisulphite sequencing will give base-pair resolution it can enable high depth counts of individual CpG sites, possibly affording the detection of small methylation changes, that might be hard to detect otherwise. This approach will allow the simultaneous testing of 96 samples (cases and controls) in one experiment along with the exploration for SNPs and cis-methylation QTLs from the same sequencing data.

5.2 Aims

I. To prepare a 96 sample (case-control) barcoded bisulphite PCR sequencing library using the Fluidigm Access Array 48.48 microfluidics system and to assess its application to parallel sequencing on the Illumina platform.

II. To measure base-pair resolution CpG methylation across the APP gene region in a AD pathology affected brain area of post-mortem tissue and to examine whether there are any differences in methylation levels that associate with disease status.

III. To test if any identified DMRs are associated with total APP protein levels in these samples.

IV. To use SNP calls from the same bisulphite sequencing data and explore the association of SNPs with nearby CpG methylation levels that may represent cis-methylation QTLs.

5.3 Samples

All samples were taken from a well-characterised cohort of brains donated to the Medical Research Council London Neurodegenerative Diseases Brain Bank. Frozen brain tissue was collected from the Superior Temporal Gyrus area (STG) reflecting an AD pathology affected tissue. DNA was extracted using a standard phenol-chloroform extraction (section 2.1). Samples consisted of both AD brains and cognitively normal controls (N=94), of which 66 are confirmed AD patients. Diagnosis was autopsy confirmed by an experienced pathologist. AD brain samples were matched with normal controls on sex and age where possible, though this was limited by the small number of normal control brains available for analysis. The mean age at death was 86.3 years (Min-Max: 69-105) for AD cases and 79.3 years (Min-Max: 40-97) for controls. Forty-one patients were female (62%) and 25 males (38%) while in the control group, 17 subjects were females (60%) and 11 subjects were males (40%). A full list of the samples used in this chapter along with demographic details is listed in Appendix A.

All samples had additional neuropathological data including Braak staging (Braak & Braak 1991) and amyloid-β scores. Braak staging uses a six-point scale (I-VI) to primarily reflect the distribution of neurofibrillary tangles (NFT) within the brain. Braak stages I and II are used when neurofibrillary tangle involvement is confined mainly to the transentorhinal region of the brain, stages III and IV when there is also involvement of limbic regions such as the hippocampus, and V and VI when there is extensive neocortical involvement (Braak & Braak 1991). Amyloid-β (Aβ) levels in the brain had previously been scored 0 to 3 and were based on pathology report immunohistochemistry results. These represent a gross measure of Aβ in the brain and are not brain area specific. In addition to the two AD pathology markers, total amyloid precursor protein (APP) levels specific to the STG brain area were also available. These had been previously determined by Western blotting using a near infrared scanner and densitometric quantification with normalization to actin levels within lane and levels expressed as arbitrary units of immunoreactivity to 22C11, a well characterized anti-human APP antibody. 22C11 binds an epitope between amino acids 66-81, within the N-terminal region of APP protein derived from codons within exons 2 and 3, which are not subject to alternate splicing. Thus the antibody recognises mature and immature forms of the holo-protein forms of all three major isoforms, APP695, APP751 and APP770, and additionally the secreted alpha and beta (sAPPα/sAPPβ) fragments derived from each by α- and β-secretase proteolytic process, respectively.

5.4 Statistical analysis

In general linear regression models were used to test the association of DNA methylation with, demographic data and AD status or disease markers Braak staging and Aβ levels or APP protein levels. In the first instance the separate relationships of age and sex alone on average ROI amplicon DNA methylation levels were tested in unadjusted linear regression. Secondly methylation at sequence features and a selection of potential DMRs were sought for independent associations with AD status by linear regression, controlling for sex, age and ApoE genotype. Where these associations presented a p-value of <0.05 APP protein levels were tested for an association with the given methylation sites. Lastly methQTL analysis was carried out with linear regressions using methylation level (%) as the outcome variable and SNP genotype values: 0 = homozygous for the reference allele; 1 = heterozygous; and 2 = heterozygous for the alternative allele (along with covariates age and sex) as the predictor variable. Where these associations presented a significant p-value (Bonferroni corrected) of <0.05, AD status was added and tested as a further predictor variable.

5.5 APP gene region methylation content assayed

5.5.1 Choice of APP regions of Interest (ROI) sequences

A number of different sequences along with the APP promoter CGI and CGI shore were chosen to investigate possible variations in CpG methylation. These further sequences included, all exons with neighbouring intronic sequence, the 5' region, 3'UTR and other intragenic sequences showing possible regulatory or enhancer potential. Alongside these we wished to further investigate the methylation status of SINE sequences in this region. Though sequencing of repeat elements is challenging, it would be interesting to see if any of the trends observed for Alu global methylation in chapter 3 are present at the APP locus. In total 144 region of interest (ROI) PCR amplicons covering 22.4 kb of sequence were chosen to be sequenced. The final APP ROI amplicons are depicted in Figure 5.6 and grouped by type in Table 5.1.

Using data from both the UCSC genome browser (human genome build hg18) and the ENCODE project, a number of different tracks were intersected using the web-based bioinformatics tools at Galaxy (Giardine et al. 2005) to reveal sequence overlaps that possibly correspond to regulatory features. While these data tracks complement each other, evidence from multiple tracks at a given sequence can be strongly suggestive of a regulatory function (The ENCODE Project, 2011). Broadly three groupings of various feature tracks were intersected resulting in

three core intersects: (1) enhancer & promoter histone modifications, (2) conservation & regulatory potential and (3) open chromatin & DNase sensitive sites. These three core intersects and their respective composite tracks are described/illustrated below (Figures 5.3-5.5).

The resulting three core intersect tracks were then used as a guide to highlight relevant sequences which were assigned to two main groups based on the type and number of data tracks they overlapped with. For instance sequences showing an overlap of intersect 1 (enhancer and promoter histone modifications) and a sub-track from intersect group 2 (conservation & regulatory potential) were assigned to the category of enhancer-conserved-regulatory potential. Six, such loci were identified and are referred to as enhancer 1 to 6 in results section. Whereas sequences showing either 2 or 3 overlaps between the sub-tracks of intersect group 3 (open chromatin & DNase sensitive sites) were assigned to the category of, open chromatin. Ten, such loci were identified and referred to as open chromatin 1 to 10 in results.

Intersect Group 1 - Enhancer & Promoter Histone Modifications:

Modifications such as methylation and acetylation to the histone proteins present in chromatin are known to influence gene expression by changing how accessible the chromatin is to transcription. These tracks show the levels of enrichment of histone marks as determined by ChIP-seq (chromatin immunoprecipitation followed by sequencing) assays. The enhancer H3K4Me1 (histone H3, Lysine 4 monomethylation) and enhancer H3K27Ac (histone H3, Lysine 2 acetylation) tracks show where modification of histone proteins is suggestive of enhancer and, to a lesser extent, promoter activity. Both H3K4Me1 and H3K27Ac are also associated with enhancers in regions downstream of transcription start sites. Whereas the promoter H3K4Me3 (histone H3, Lysine 4 trimethylation) track shows a histone mark associated with promoters. Though histone modifications, especially H3K4Me1, are quite broad and the actual enhancers are typically just a small segment of the area marked by these histone modifications, they provide an indication as to the location of potential regulatory elements. Also while there are data for number of cell lines available, we only looked at the data from human epidermal keratinocytes (NHEK) as this represented the closest matching tissue to brain, as both originate from the ectoderm in the early embryo. Also the UCSF brain DNA methylation track was included in this intersect as a reference to show regions of DNA methylation in brain.

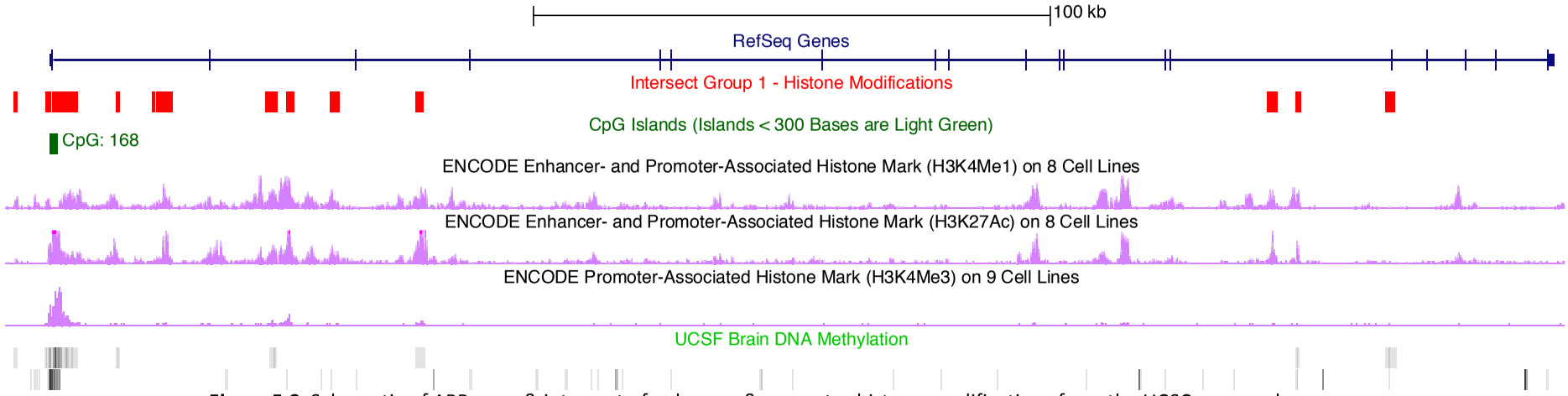


Figure 5.3 Schematic of APP gene & intersect of enhancer & promoter histone modifications from the UCSC genome browser [red track: intersect of the data from the composite tracks]

Intersect Group 2 - Conservation & Regulatory Potential:

Evolutionary conservation over a sequence is an important indicator of biological function, including regulatory potential. The vertebrate basewise conservation track shows peaks from multiple alignments of 44 vertebrate species and measurements of evolutionary conservation (using the PhyloP method), whereas the ESPERR regulatory potential track displays regulatory potential scores computed from alignments of human, chimpanzee, macaque, mouse, rat, dog and cow. These sequence conservation tracks are then complemented with data from the HMR conserved transcription factor binding sites track (location and score of transcription factor binding sites conserved in the human/mouse/rat alignment) and actual experimental data from the ENCODE transcription factor track which shows transcription factors that bind to DNA as supported by ChIP-seq experiments. It was hoped that combining these tracks would allow the identification of highly conserved regulatory sequences, which could then be combined with the data from intersect 1 to yield regions with conserved enhancer potential outside of the core APP promoter region.

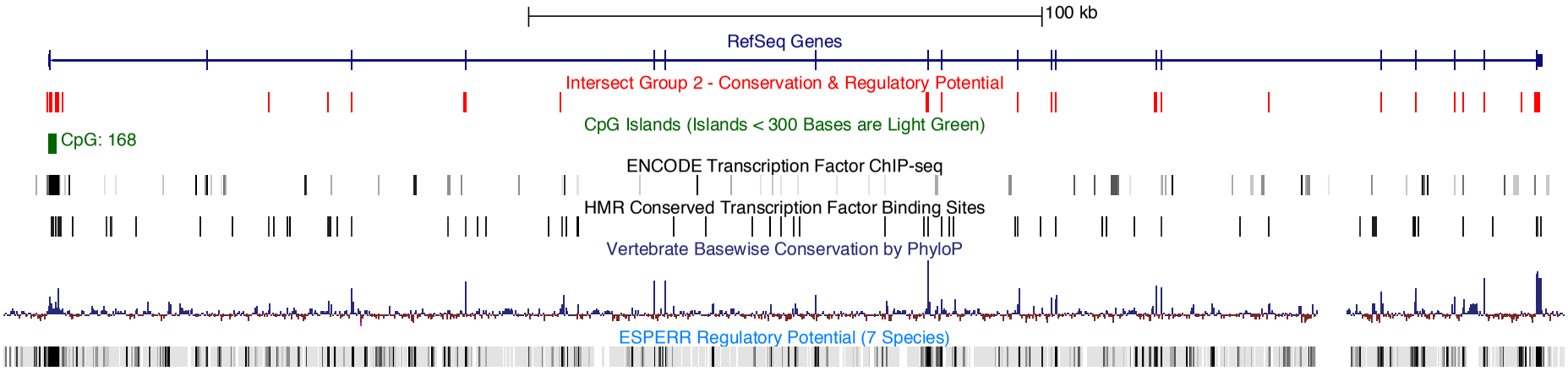


Figure 5.4 Schematic of APP gene & intersect of conservation/regulatory potential from the UCSC genome browser [red track: intersect of the data from the composite tracks]

Intersect Group 3 - Open Chromatin & DNase Sensitive Sites:

All the data tracks used in this intersect group represent accessible/open chromatin sequences mainly marked by DNase hypersensitive sites (DHSs). Regulatory regions in general tend to be DNase hypersensitive owing to their remodelled nucleosome state; within the cluster of DHSs a regulatory factor binding the DNA can also protect the underlying sequence from cleavage by DNase, leaving a nucleotide-resolution footprint. Though this highlights sequences with regulatory potential, at the functional level this provides little information beyond that. The ENCODE digital DNase I hypersensitivity clusters and University of Washington digital DNase genomic footprinting, show data on DHSs and DNase footprints respectively. While the Open Chromatin, Duke/UNC/UTs track shows data from three types of experiment, DNase I hypersensitivity, Formaldehyde-Assisted Isolation of Regulatory Elements (FAIRE) to identify nucleosome-depleted regions, and ChIP-seq for select regulatory factors. The data from ChIP experiments help in the functional annotation of the open chromatin regions identified by DHS mapping and FAIRE. Again these data tracks present data for numerous cell lines and NHEK cells were chosen as a surrogate for brain tissue.

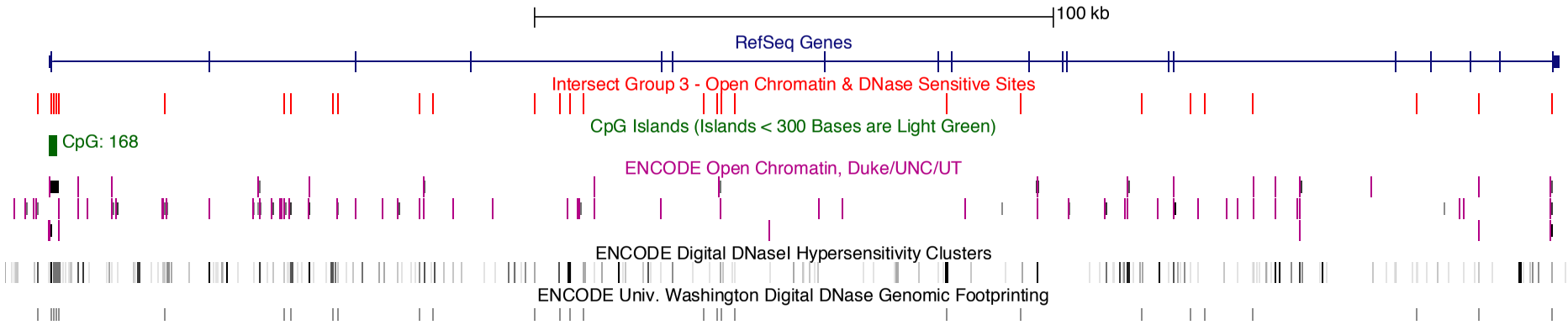


Figure 5.5 Schematic of APP gene & intersect of open chromatin and DNaseI clusters from the UCSC genome browser [red track: intersect of the data from the composite tracks]

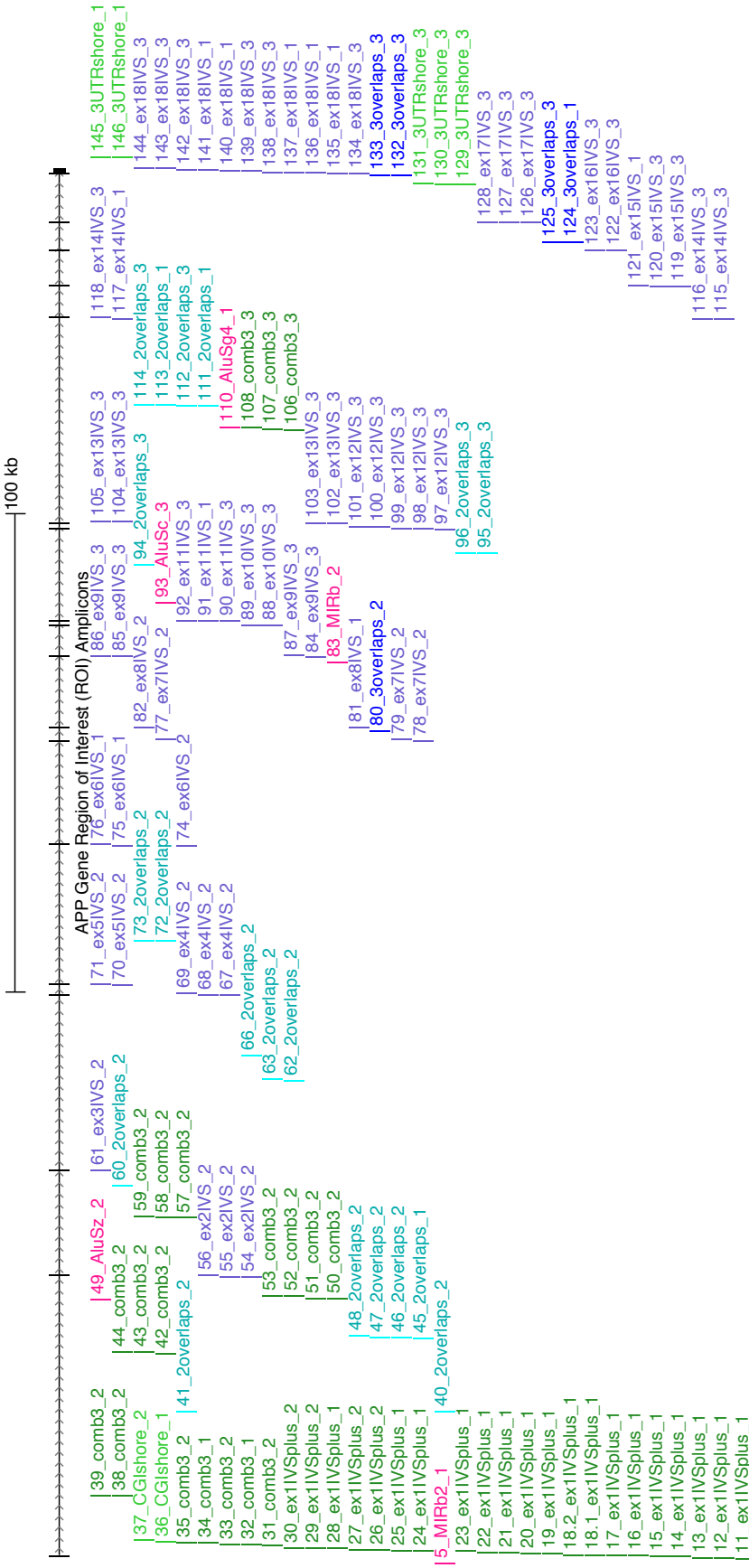


Figure 5.6 Schematic of APP gene region & location of ROI amplicons

Feature type	Description	Number of amplicons
CGI and exon 1	Amplicons spanning the entire CGI which also covers exon1	26
CGI shores	Sequences ± 2kb from CGI	4
Exons 2-18	All transcribed exons and partial intronic sequences	58
Enhancer-conserved-regulatory potential	Overlap of intersect 1 and a subtrack from intersect group 2 [marked as comb3 in figure 5.6]	20
Open chromatin-DNase sensitivity	Either 3 or 2 overlaps of subtracks from intersect group 3 [marked as 3overlaps and 2overlaps in figure 5.6]	24
3'UTR	Sequence downstream from exon 18	5
SINE element	Alu and MIR repeat elements ~300bp in size	7

Table 5.1 Summary of amplicon numbers by feature type

5.5.2 Validation and optimisation of targeted bisulphite PCR/sequencing primers

As Fluidigm 48.48 Access Array primary PCR cycling conditions incorporate what is known as C_0T cycling and require a relatively low final concentration of target specific primers. All of the primer pairs were validated to ensure they met with these platform specific conditions. In addition, further optimisation of the PCR was required to achieve a standardised annealing temperature across all primer pairs. Initial optimisation showed that most of the target-specific primers (primary PCR) worked at the standardised annealing temperature of 60^{o}C, generating a product in the expected size range of 150-250bp (with a few amplicons expected to amplify ~300bp). Though there was some very faint non-specific banding in some products this was not present in final secondary PCR products. A small number of primer pairs failed to produce a strong product band or showed the presence of some strong non-specific products, such as that seen for some product bands in the last row of the agarose gel picture in Figure 5.7.

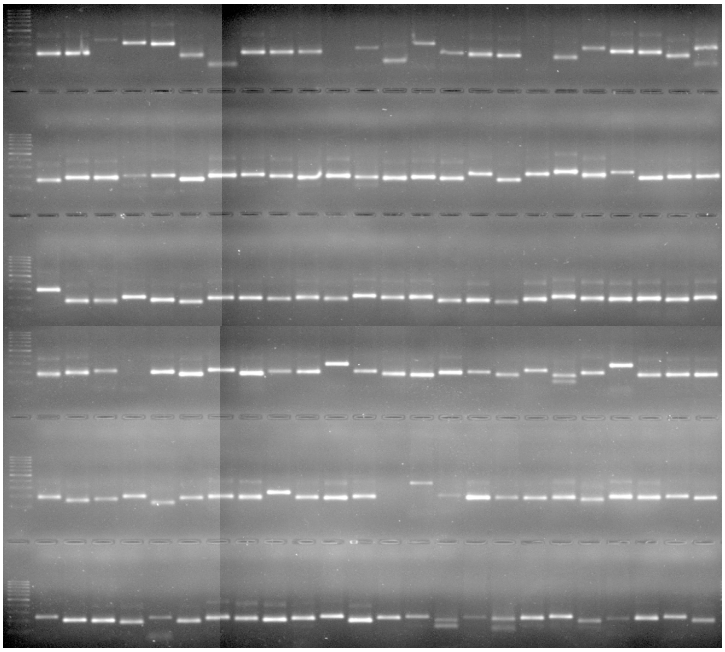


Figure 5.7 Agarose gel electrophoresis of 144 APP bisulphite PCR products
[2% agarose gel; lane on the far left: 100bp DNA molecular weight marker; 5 µL of each PCR product loaded per well]

Further optimisation of these poor performing primers was achieved by carrying out titrations of MgCl and primer concentrations (Figure 5.8). Overall it was preferable to keep the PCR reagent components the same for all PCRs and to adjust the primer concentrations; as a result most of the primers that need further optimisation were generally adjusted for primer concentration in the PCR reaction, while some persistently difficult primers were re-designed and re-tested (Figure 5.9).

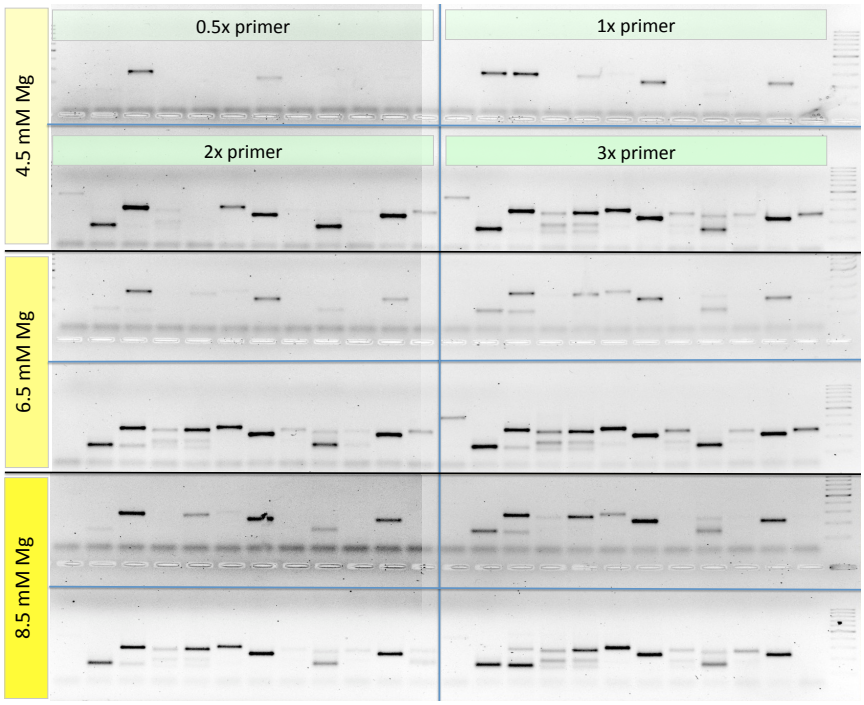


Figure 5.8 Agarose gel electrophoresis of MgCl and primer conc. titrations of APP bisulphite PCR products [2% agarose gel; lane on the far right: 100bp DNA molecular weight marker; 5 µL of each PCR product loaded per well; four different primer concentrations across three different MgCl (Mg) concentrations were tested to ascertain which standard MgCl concentration worked best for all the samples tested, while optimal primer concentrations were noted per individual primer set tested]

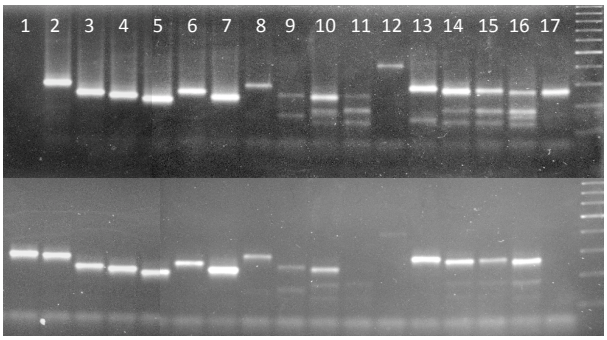


Figure 5.9 Agarose gel electrophoresis of APP bisulphite PCR optimisations [2% agarose gel; lane on the far right: 100bp DNA molecular weight marker; 5 µL of each PCR product loaded per well; upper row/gel: un-optimised PCRs; lower row/gel: PCR products after optimisation; lanes 1-8: redesigned primer pairs; lanes 9-17 primer pairs requiring optimisation]

5.5.3 Fluidigm 48.48 Access Array PCR based Bisulphite Sequencing Library

On completion of the 48.48 Access Array PCR and the secondary barcode-adaptor incorporation PCR, the PCR products from a small number of individual samples from each array were checked on an Agilent 2100 Bioanalyzer DNA chip before post-PCR purification and then again for each set of pooled and cleaned 48 samples from each array. Bioanalyser profiles from a single sample pre-clean up and the respective pool from that array for the three different 48 primer sets post-clean up are shown in Figure 5.10. The profiles for each primer set were in the expected size ranges, indicating successful PCR library generation.

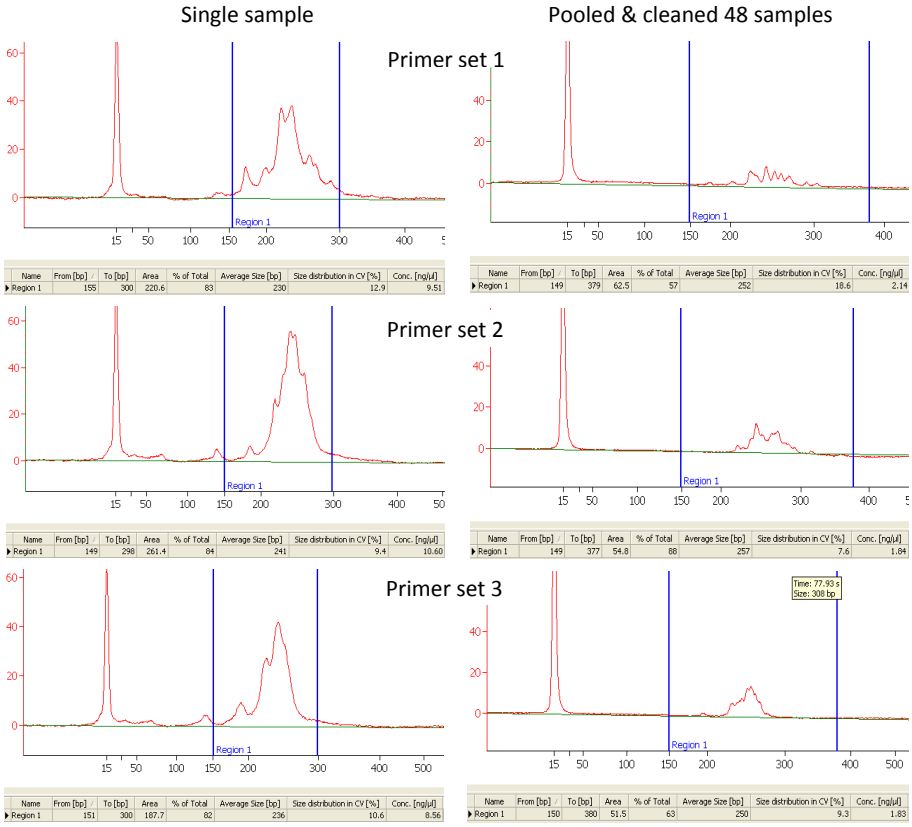


Figure 5.10 Screenshots of Agilent bioanalyzer profiles of Access Array PCR products [column on left: profiles for a single sample from each 48 primer set; column on right: profile from pooled and cleaned products from a single Access Array comprising 48 primer and 48 samples; x scale: bp; y scale: RFU]

5.5.4 Sequencing quality control as indicated by FastQC

Before analysing the sequencing data some simple quality control checks were performed to ensure that the raw data from individual samples was of good quality and that there were no problems or biases in the generated data. The FastQC reports were used to identify problems, which may have originated either in the sequencer or in the starting library material. The FastQC report statistics generated for sample 48 read 1 are shown in Figures 5.11-5.16. The sequence length distribution of the read is as expected with a length of 101bp (Figure 5.11) and average sequence quality Phred score is 36 across the read (Figure 5.12). The pre base sequence quality score are also as expected as seen in Figure 5.13, with a slight drop off at the end of the read and a very slight base call of N at base position 9 (Figure 5.14). While the per base sequence content shows an elevated proportion of thymine and a reduced cytosine (Figure 5.15) along with a reduced GC content (Figure 5.16), this is as expected given the data are bisulphite treated DNA sequencing.

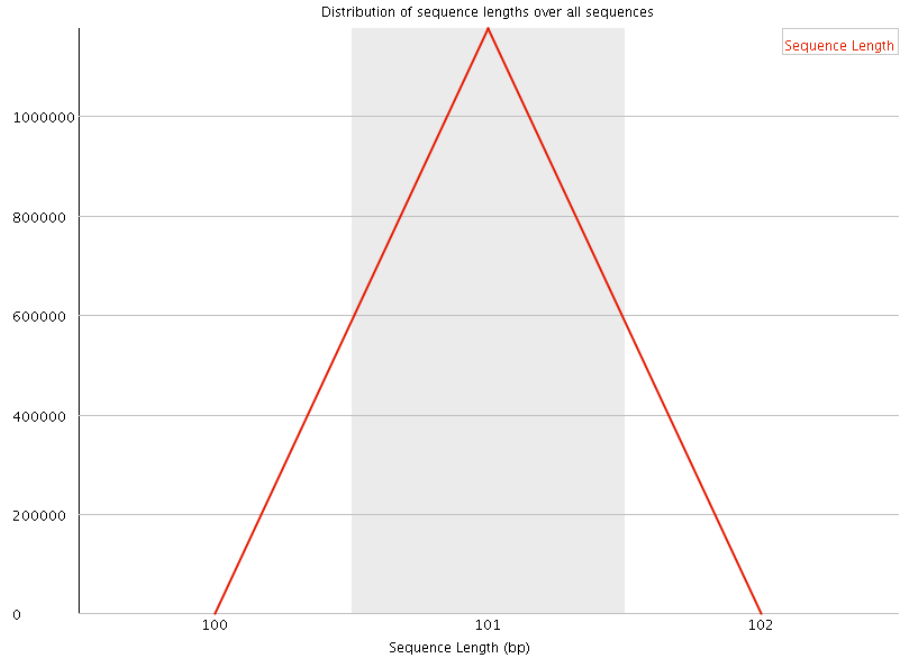


Figure 5.11 FastQC report - Sequence Length Distribution of sample 48 read 1

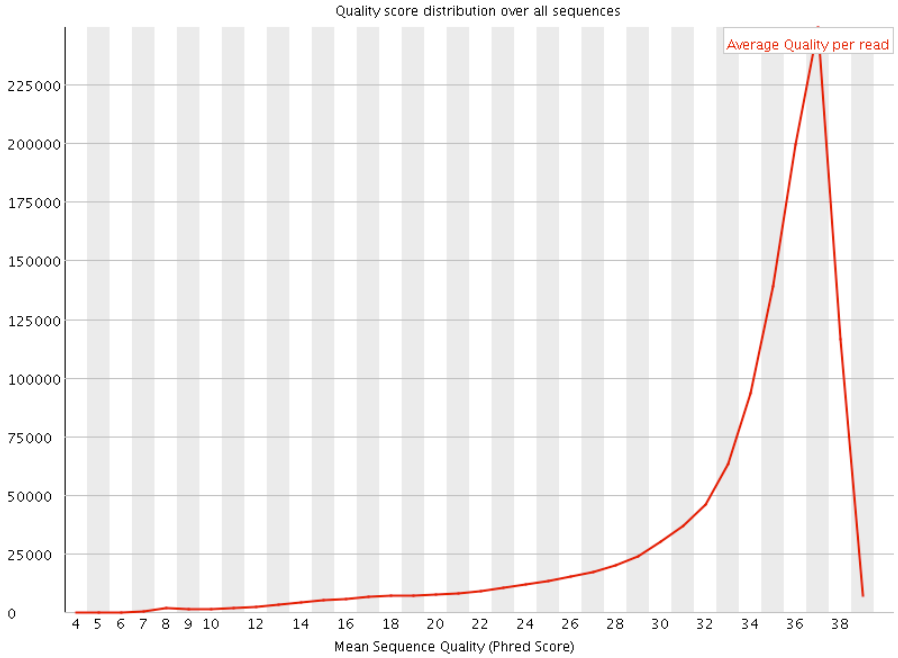


Figure 5.12 FastQC report - per sequence quality scores of sample 48 read 1

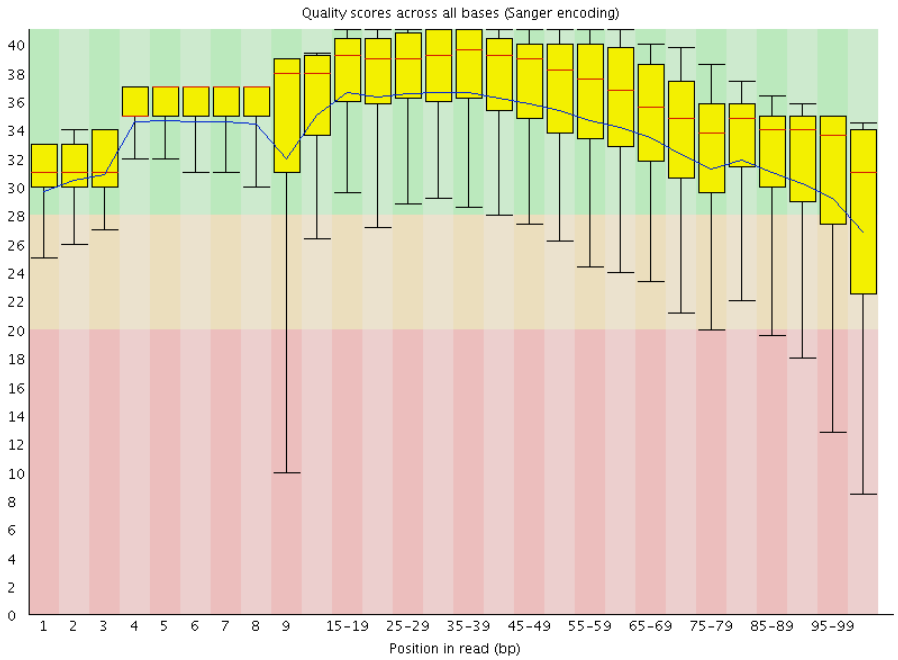


Figure 5.13 FastQC report - per base sequence quality of sample 48 read 1

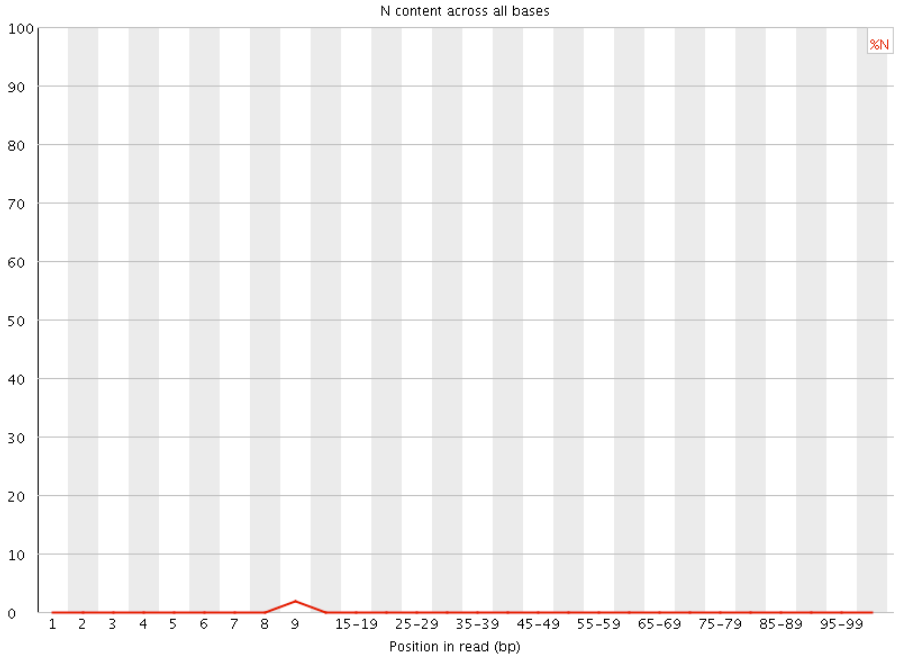


Figure 5.14 FastQC report - per base N content of sample 48 read 1

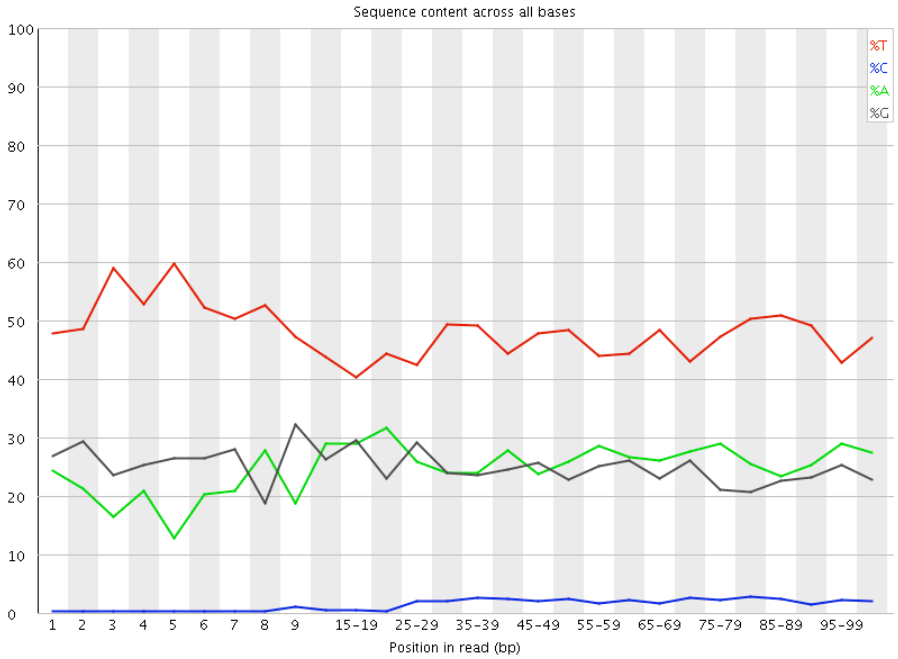


Figure 5.15 FastQC report - per base sequence content of sample 48 read 1

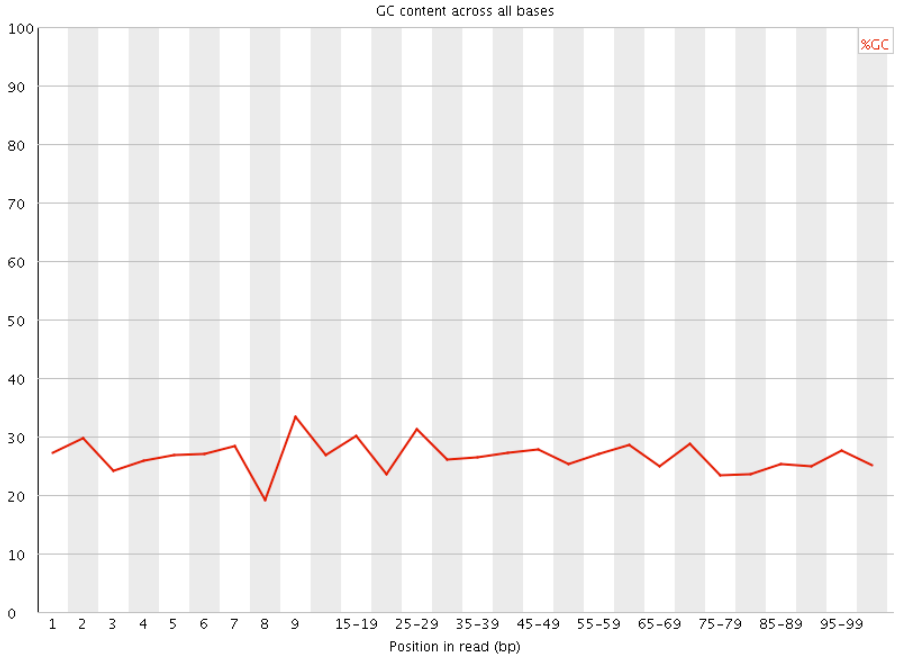


Figure 5.16 FastQC report - per base GC content of sample 48 read 1

5.6 Results

5.6.1 Bisulphite sequencing alignment statistics

Individual sample data for the 94 case control samples plus fully methylated and unmethylated control samples were aligned to the human genome (hg18) using the Bismark bisulphite sequencing alignment package (Krueger & Andrews 2011). Based on statistics derived from the alignments, in total the data generated from the single lane of Illumina HiSeq 2000 was 131323730 paired-end (PE) reads with 87317749 PE reads mapping correctly to the genome, resulting in a mapping efficiency of 66% (Figure 5.17 and Table 5.2).

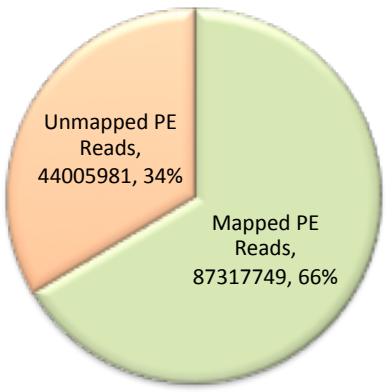


Figure 5.17 Pie chart of PE reads sequencing alignment

Description	Total	Per sample average
Analysed PE reads	131323730	1367955
Mapped PE reads	87317749	909559
Unmapped PE reads	44005981	458395
Mapping efficiency (%)	66%	66%

Table 5.2 Sequencing alignment summary statistics

Bismark also provided some preliminary statistics on the gross CpG methylation levels from the aligned reads. These preliminary statistics indicate that gross average methylation; in AD samples is 32.6% and 34.8% in controls. Also the fully methylated control shows a gross CpG methylation of 93.3% and the unmethylated control sample shows a 1.6% level of gross methylation (Table 5.3). These methylation values for the fully methylated and unmethylated control samples are in the expected range, as neither is actually fully 100% or 0% methylated.

Measure at CpG sites	AD samples average	Control samples average	Methylated control sample	Unmethylated control sample
Total methylated C's in CpG context	1551655	1322299	4053082	103264
Total C to T conversions in CpG context	3231231	2627722	292312	6286457
C methylated in CpG context (%)	32.6%	34.8%	93.3%	1.6%

Table 5.3 Gross levels of CpG methylation as summarised by Bismark alignment

5.6.2 Quantification of CpG methylation

Accurate quantification of CpG sites across the APP gene data and within ROI amplicons for the 94 test samples was achieved using the SeqMonk package (http://www.bioinformatics.bbsrc.ac.uk). Further QC measures were used to remove samples and amplicons presenting poor data. Based on low individual sample correlations (r=<0.65) in CpG methylation calls against all other samples, 15 samples were removed as poor data, leaving 79 test samples for all further analysis (AD=58; Ctrl=21). In addition from the original 144 amplicons, 114 ROIs were identified that presented good methylation call data with a minimum read depth of 200 reads across most of the test samples and only these were carried over for further analysis. The total number of individual CpG sites that were covered by the 114 ROI sequences was 335. Methylation levels were calculated for both the individual 335 CpGs and the 114 ROI (average methylation of all the CpGs within a given ROI sequence). In the following sections data corresponding to a ROI sequence is identified by a number with a ‘f’ prefix, whereas a number prefixed with a ‘c’ identifies data for a single CpG.

PCR amplicons of particular interest that either dropped out or failed to produce consistent data across all samples tested included five ROI amplicons that were in the CGI (Figure 5.18) and three others which covered exon 6, 10 and 15. Consequently these eight key ROIs were excluded from further analysis.

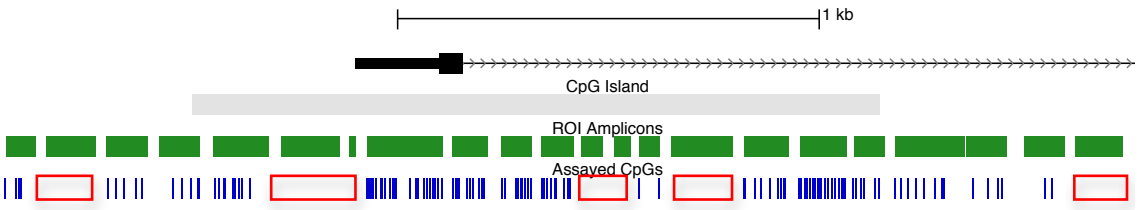


Figure 5.18 Schematic of APP CGI region ROI amplicons and successfully assayed CpGs [grey: CGI; green: ROI amplicons; blue vertical bars: successfully assayed CpGs; red box: dropped out ROI amplicons]

5.6.3 Age and sex associations with average ROI methylation

Unadjusted linear regression of age and sex separately with the average ROI methylation levels from each of the 114 ROIs indicated there to be significant associations at certain ROIs, as summarised in Tables 5.4 & 5.5 respectively. Specifically males presented with less methylation than females at five separate ROI spanning the entire gene from IVS1 (f45 and f68), IVS2, IVS13 through to exon 18 (f28) where the significant difference in average methylation was greatest at ~8.7% (Figure 5.19). Though most of the seven ROIs showing a marginally significant association of age with average methylation levels at the particular ROI, present with an increase in methylation with age, two of these ROIs in IVS3 (f79 and f81) show the inverse relationship with age (Table 5.5). All further analyses were carried out using age and sex along with ApoE genotype as a covariate.

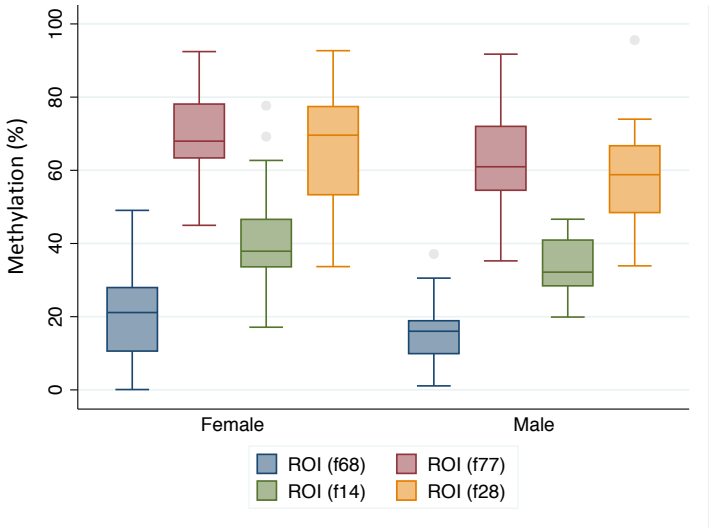


Figure 5.19 Box plot of average ROI methylation by sex

Methylation ROI	Coefficient	Std. Err.	t	*p*	95% CI	
IVS1 (f45)	-0.303	0.112	-2.71	0.008	-0.525	-0.080
IVS1_Enhancer 4 (f68)	-5.364	2.346	-2.29	0.025	-10.036	-0.692
IVS2_Open Chromatin 3 (f77)	-6.398	2.899	-2.21	0.030	-12.172	-0.625
IVS13 (f14)	-5.529	2.502	-2.21	0.030	-10.512	-0.547
Exon 18 (f28)	-8.694	3.602	-2.41	0.018	-15.871	-1.516

Table 5.4 Significant associations of sex with average amplicon methylation

Methylation ROI	Coefficient	Std. Err.	t	*p*	95% CI	
IVS3 (f79)	-0.177	0.088	-2.03	0.046	-0.352	-0.003
IVS3_Open Chromatin 4 (f81)	-0.135	0.060	-2.25	0.027	-0.255	-0.016
IVS3_Open Chromatin 4 (f82)	0.116	0.052	2.25	0.027	0.013	0.219
Exon 5 (f86)	0.168	0.071	2.37	0.020	0.027	0.309
Exon 11 (f102)	0.571	0.188	3.03	0.004	0.193	0.948
IVS11 (f106)	0.227	0.084	2.69	0.009	0.058	0.395
Exon 17 (f23)	0.170	0.080	2.12	0.040	0.008	0.332

Table 5.5 Significant associations of age with average amplicon methylation

5.6.4 CpG methylation across ROI features

Methylation levels from the 114 different ROI were grouped according to feature type and average methylation levels calculated for these features. As shown in Figure 5.20, the average methylation levels at each feature type vary across the gene and as expected the lowest levels are at the CGI (~0-3%) and the highest levels at Alu repeats (~95%) and the 3'UTR (~92%) sequences. Features with methylation in the range of 20-50% include the 5' region, the CGI shore (3' to the CGI), enhancers and open chromatin, and while this range also includes the SINE repeat class sequence of MIRb, this feature was only covered by data from a single CpG site and thus unlikely to be a accurate representation of MIRb sequences. Further to this, of the seven SINE ROI sequences that were selected for sequencing only two of these actually produced consistent amplification resulting in good methylation calls, one being a Alu sequence in intron 1 and the MIRb sequence as already mentioned in intron 8. Average methylation levels at the CGI shore (5' to the CGI), exons and introns (IVS) fall in the range of 70-80%. In comparing the average methylation levels at features type by AD status there are no striking differences at any particular feature type, excluding the MIRb data as noted above.

Adjusted linear regressions (controlling for age sex and ApoE as noted above) were run for each of the feature types against AD status to test for any significant associations. However no significant associations were obtained for these analyses and for which β-coefficient and p-values are summarised in Table 5.6.

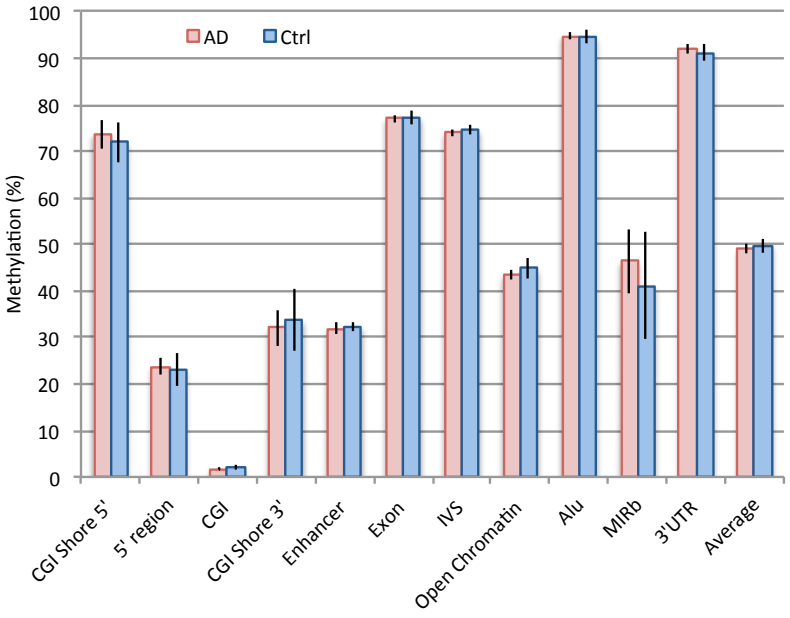


Figure 5.20 Bar chart of average CpG methylation by feature type and AD status [whiskers: 95% CI]

Methylation Feature	Coefficient	Std. Err.	t	P-value	95% CI	
CGI Shore 5'	0.788	3.19	0.25	0.806	-5.574	7.149
5' region	0.116	2.064	0.06	0.955	-3.997	4.23
CGI	-0.441	0.287	-1.53	0.129	-1.012	0.131
CGI Shore 3'	-2.229	4.271	-0.52	0.603	-10.74	6.281
Enhancer	-0.124	1.138	-0.11	0.914	-2.391	2.143
Exon	-0.156	0.966	-0.16	0.872	-2.08	1.768
IVS	-1.185	0.812	-1.46	0.149	-2.802	0.433
Open Chromatin	-0.784	1.237	-0.63	0.528	-3.248	1.681
Alu	-0.12	0.843	-0.14	0.887	-1.801	1.561
MIRb	1.964	7.52	0.26	0.795	-13.055	16.983
3'UTR	0.421	1.204	0.35	0.728	-1.977	2.819
Average across all amplicons	-0.573	1.057	-0.54	0.589	-2.68	1.533

Table 5.6 Summarised β-coefficient & p-value for tested association of AD status with methylation features

5.6.5 Methylation profile across exons and introns

Methylation levels from CpG sites that correspond to individual exons and introns (IVSs) of the APP gene were averaged and plot by AD status (Figure 5.20). As a general trend the exons show higher levels of methylation relative to the introns (as already indicated in Figure 5.20). Visual inspection of the plot shows there to be two potential differentially methylated regions (DMRs) marked by the arrows around, exon 5 and exon/intron 11.

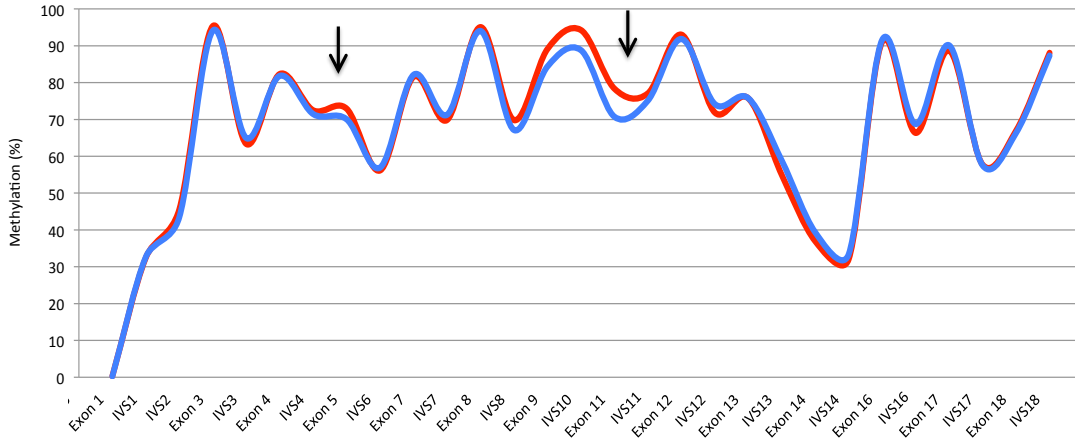


Figure 5.21 Smoothed scatter plot of average methylation levels across exons and introns by AD status [averaged data from all CpG sites covering each specified exon and IVS region; red line: AD samples; blue line: control samples; black arrows: potential DMRs]

5.6.6 Methylation profile across the 335 CpG sites

An overview of APP methylation as ascertained by all the bisulphite sequencing data is represented in a profile plot of the assayed 335 CpG sites in Figure 5.22 (by AD status). From around 75-90% methylation levels in the 5' region levels drop to ~0.5% in the CGI and then increase back to 20-50% in intron 1, demonstrating the signature pattern of a hypomethylated gene promoter/CGI. At the gross methylation level there are virtually no consistently differentially methylated sections of the CGI between AD and control. Though upon closer inspection of the scaled CGI methylation data in the lower plot of CGI methylation in Figure 5.22, there are a couple of spikes of higher methylation for the AD samples compared to the controls at the start of the CGI and a few spikes of higher methylation in controls compared to AD samples in the middle of the CGI, these may be due to poor data at these individual CpG sites.

Further profiles of the assayed 335 CpG sites by dichotomised Braak staging (Figure 5.23) and gross brain Aβ levels (Figure 5.24) show similar patterns to that generated by AD status (Figure 5.22). Across the three profiles (Figures 5.22-5.24) there are four potential DMRs where two or more neighbouring assayed CpG sites consistently show a similar pattern. The arrows on Figure 5.22 mark these four sites with in the 5'region, IVS1, exon 5 and exon 11/IVS11. These latter two potential DMRs were also highlighted by the exon and intron profile by AD status (Figure 5.21) described above. Given that the profiles generated by the three different measures of disease are very similar and in an attempt to limit the burden of multiple testing, only AD status was investigated for potential associations with methylation.

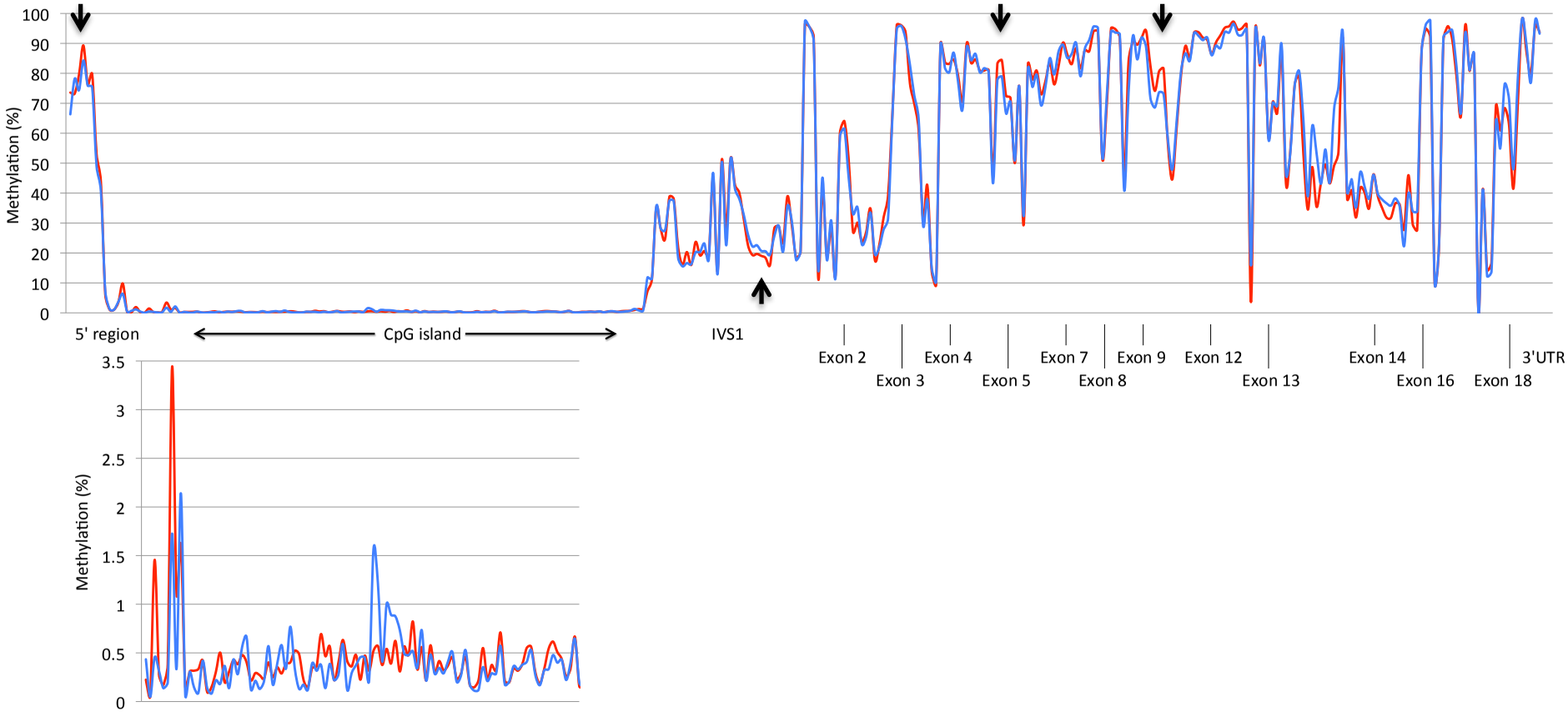


Figure 5.22 Smoothed scatter plot of average CpG methylation levels across the APP gene by AD status
[data from all assayed 355 CpG sites are shown; red line: AD samples; blue line: control samples; lower plot: scaled CGI methylation; black arrow: potential DMRs]

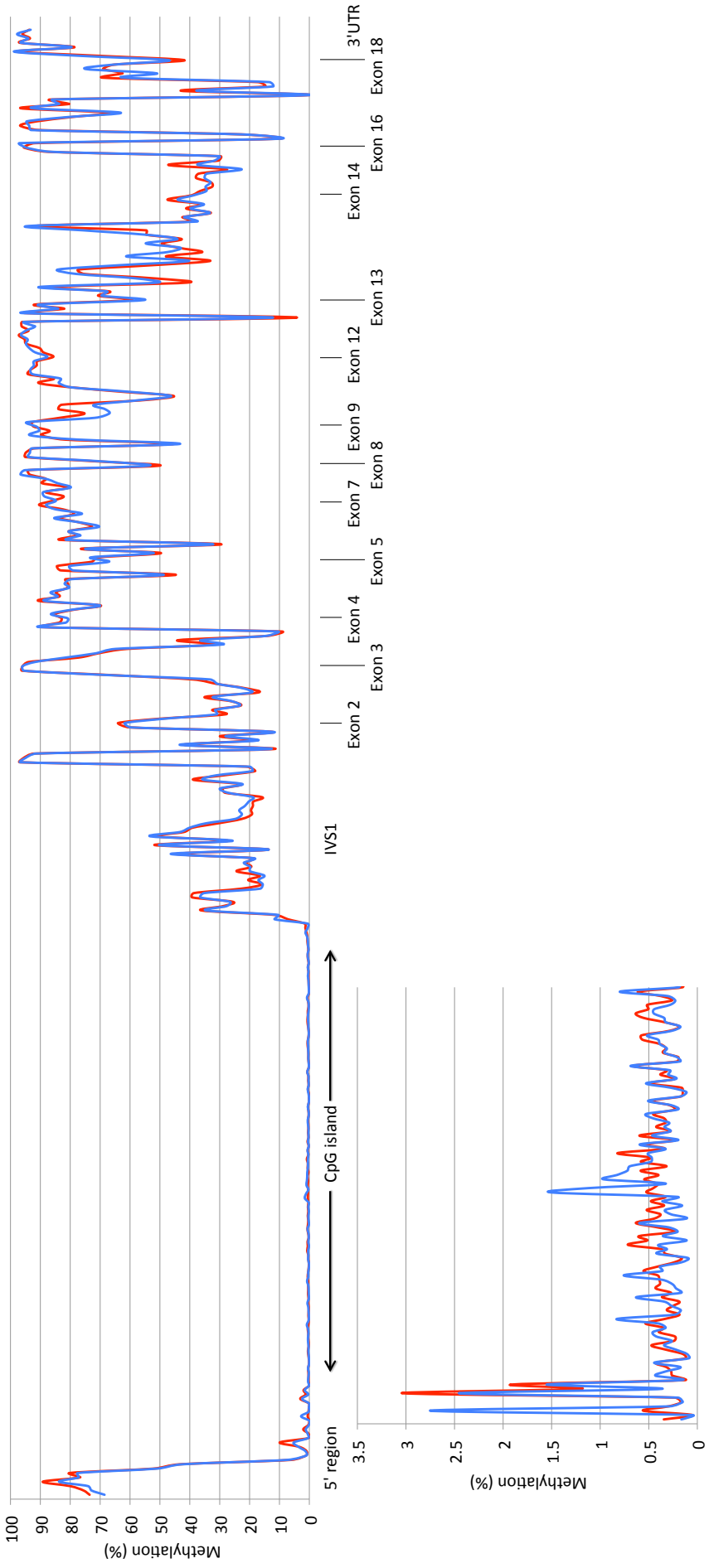


Figure 5.23 Smoothed scatter plot of average CpG methylation levels across the APP gene by grouped Braak stage
[data from all assayed 355 CpG sites are shown; red line: Braak stages 4-6 samples; blue line: Braak stages 0-3 samples; lower plot: scaled CGI methylation]

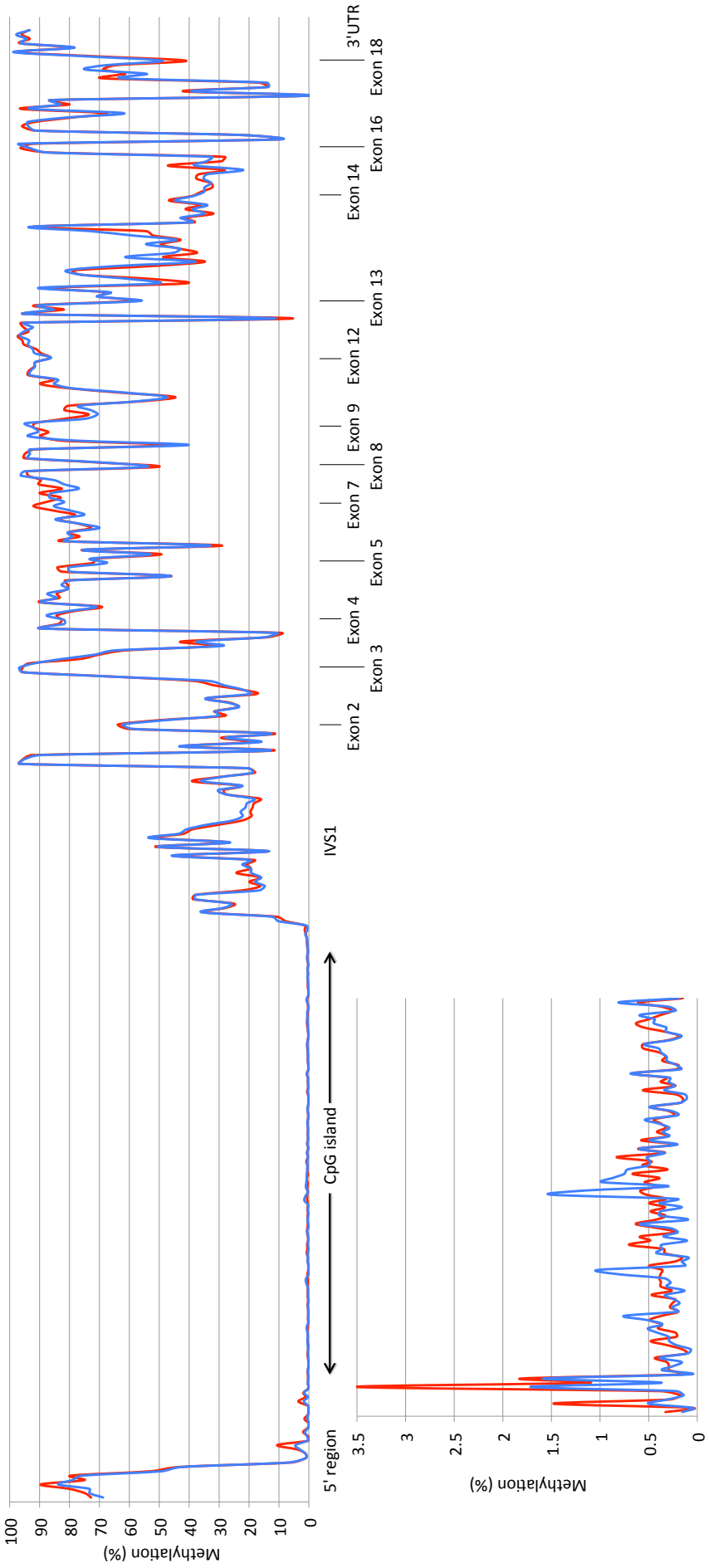


Figure 5.24 Smoothed scatter plot of average CpG methylation levels across the APP gene by grouped amyloid-β levels
[data from all assayed 355 CpG sites are shown; red line: amyloid-β levels 2-3 samples; blue line: amyloid-β levels 0-1 samples; lower plot: scaled CGI methylation]

5.6.7 Analysis of potential DMRs across the APP gene region

The four potential DMRs identified in the previous section are shown in the Figure 5.25 bar chart. Average methylation levels at the each of the three CpG sites covering each of the four DMRs exhibit hypermethylation in AD samples relative to controls, with the exception of the three CpGs covering the IVS1 enhancer 3 sequence, where AD samples are hypomethylated relative to control samples (Figure 5.25). AD status tested individually against methylation levels at these 12 CpG sites by linear regressions is summarised in Table 5.7. One CpG in the 5'region (c1) and the three CpG sites covering exon 5 (C195, c196 and c197) show a significant association of AD status with methylation levels, however these p-values do not withstand correction for multiple testing (Table 5.7).

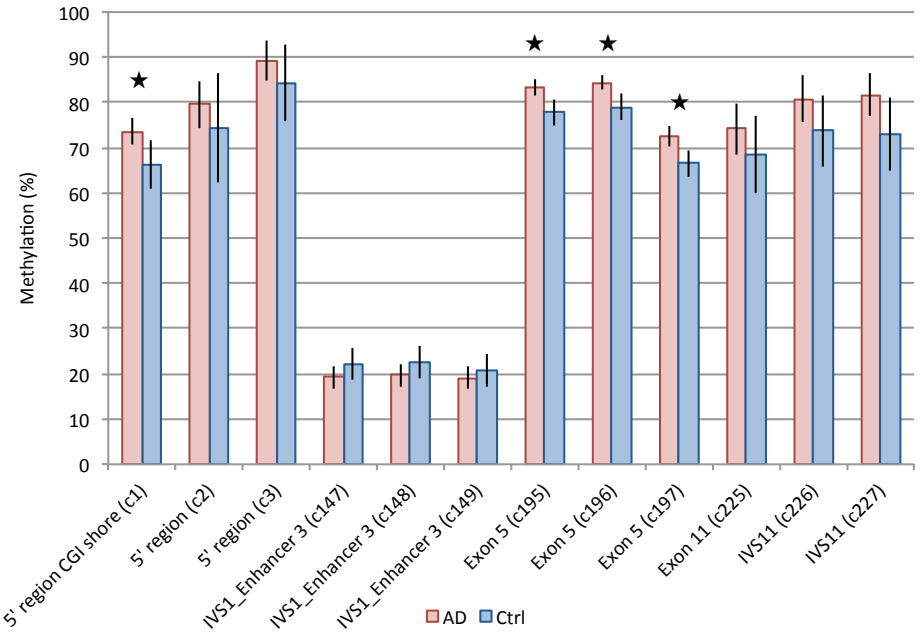


Figure 5.25 Bar chart of average CpG methylation levels at potential DMRs by AD status [potential DMR sites corresponding to Table 5.7 are shown here; whiskers: 95% CI; black star: based on uncorrected p-value AD status is significantly associated with methylation at this site]

Methylation CpG	Coefficient	Std. Err.	t	P-value	Corrected P-value	95% CI	
5' region CGI shore (c1)	7.481	3.212	2.33	0.023	0.276	1.076	13.886
5' region (c2)	7.359	6.16	1.19	0.237	2.844	-4.98	19.698
5' region (c3)	6.458	4.935	1.31	0.196	2.352	-3.428	16.344
IVS1_Enhancer 3 (c147)	-3.029	2.696	-1.12	0.265	3.180	-8.403	2.344
IVS1_Enhancer 3 (c148)	-3.086	2.714	-1.14	0.259	3.108	-8.497	2.325
IVS1_Enhancer 3 (c149)	-1.899	2.58	-0.74	0.464	5.568	-7.043	3.245
Exon 5 (c195)	5.275	1.938	2.72	0.008	0.096	1.414	9.137
Exon 5 (c196)	4.582	1.927	2.38	0.020	0.240	0.743	8.421
Exon 5 (c197)	5.582	2.312	2.41	0.018	0.216	0.975	10.19
Exon 11 (c225)	0.039	5.707	0.01	0.995	11.94	-11.413	11.49
IVS11 (c226)	1.859	5.251	0.35	0.725	8.700	-8.678	12.396
IVS11 (c227)	3.349	4.945	0.68	0.501	6.012	-6.574	13.273

Table 5.7 Summarised β-coefficient & p-value for tested association of AD status with potential DMRs [red p-value:<0.05; corrected P-vales with a Bonferroni correction]

Association of AD status and methylation site c1 within the 5' region

Three CpG sites (c1, c2 and c3) in the 5' region potential DMR and their location relative to the CpG island (CGI) is shown in Figure 5.26. One of these CpG sites c1, lies ~2kb upstream of the CGI (CGI shore) and two sites (c2 and c3) in the 5' region between the CGI shore and the CGI proper.

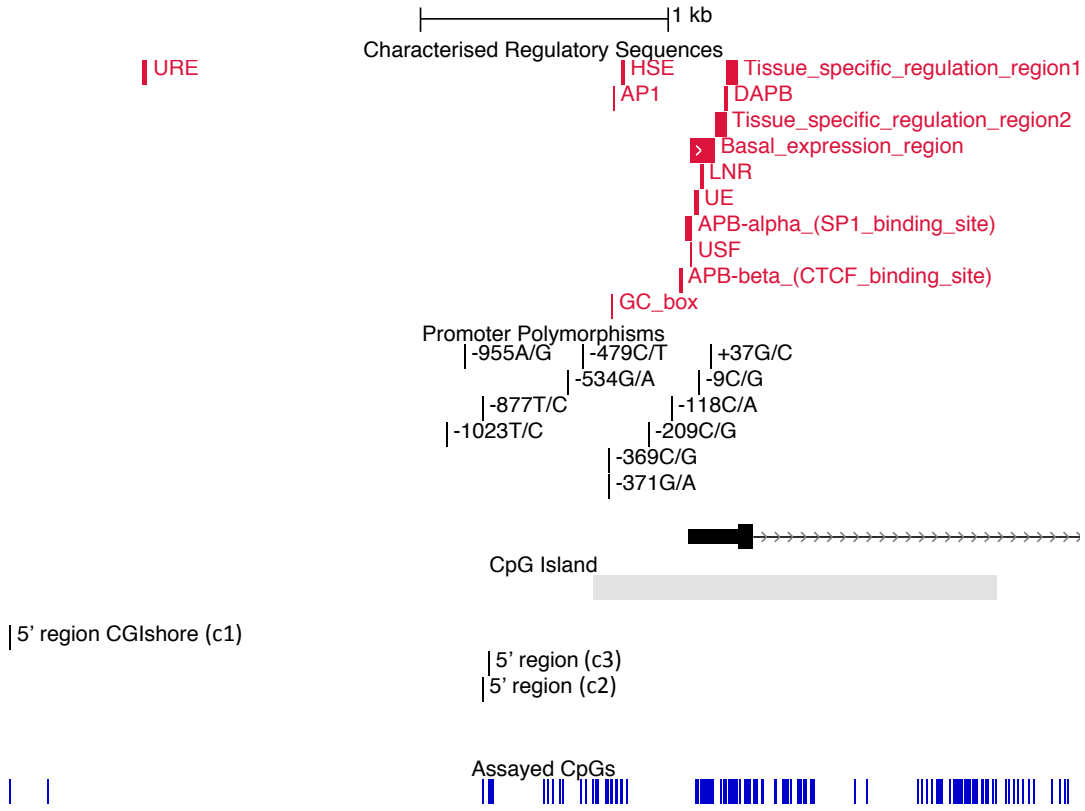


Figure 5.26 Schematic of APP gene 5'/CGI region and location of methylation sites c1 c2 & c3 [top part of figure: characterised regulatory sequences and promoter polymorphisms; middle part of figure: APP gene and CGI; bottom part or figure methylation sites c1, c2 & c3 as indicated by black vertical bars; followed by assayed CpGs indicated by blue vertical bars]

AD status is significantly associated with methylation the c1 methylation site at the 5' region CGI shore, (Table 5.7), where on average AD samples have 7.5% more methylation compared to controls ($\beta = 7.481$, SE = 3.212, $p = 0.023$), this being the largest difference between cases and controls at any of the tested potential differentially methylated CpG sites. While upon testing for an association of APP protein levels with methylation levels at the c1 CpG site, this proved negative ($\beta = -0.017$, SE = 0.038, $p = 0.668$).

Though not significant both CpG site c2 and c3 are in close proximity to the previously identified promoter polymorphisms rs466433 (-877T/C) and rs364048 (-955A/G), with CpG c2 being only

4bp upstream of -877T/C and c3 being 22bp downstream from -877T/C (Figure 5.27) and though the haplotype -877T/-955A is reported to lead to a 4-fold increase in transcription (Lv et al. 2008), there was no association of either of these SNPs with methylation at CpG c2 or c3 (see section 5.6.9). Further still all but three of the samples genotyped for these SNPs are homozygous for the reportedly risky -877T/-955A haplotype, whereas the other three samples are heterozygous for these SNPs and fall within the AD samples.

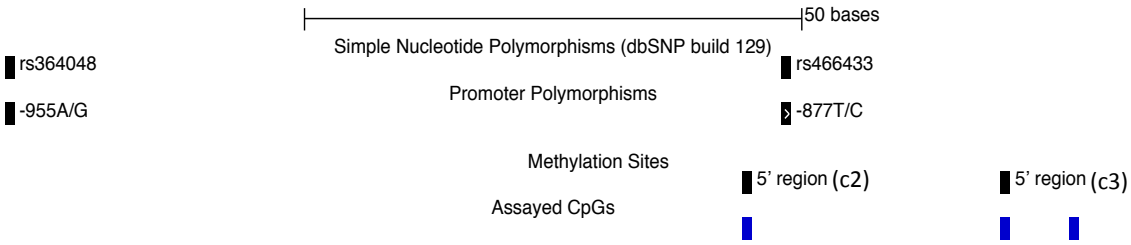


Figure 5.27 Schematic of 5' region CpGs c2 & c3 in relation to promoter polymorphisms [dbSNP build 129 track at the top followed by reported promoter polymorphisms; methylation site as indicated by black boxes; assayed CpGs indicated by blue boxes in lower part of figure]

Association of AD status at methylation sites c195, c196 and c197 across exon 5

AD status is significantly associated with methylation at three CpG sites (c195, c196 and c197) in exon 5. Where on average AD samples have ~5% more methylation compared to controls (c195 β = 5.275, SE = 1.938, p = 0.008; c196 β = 4.582, SE = 1.927, p = 0.020; and c197 β = 5.582, SE = 2.312, p = 0.018). While upon testing for an association of APP protein levels with methylation levels at the three CpG sites, proved negative (c195 β = -0.025, SE = 0.023, p = 0.287; c196 β = -0.013, SE = 0.023, p = 0.563; and c197 β = -0.011, SE = 0.027, p = 0.682).

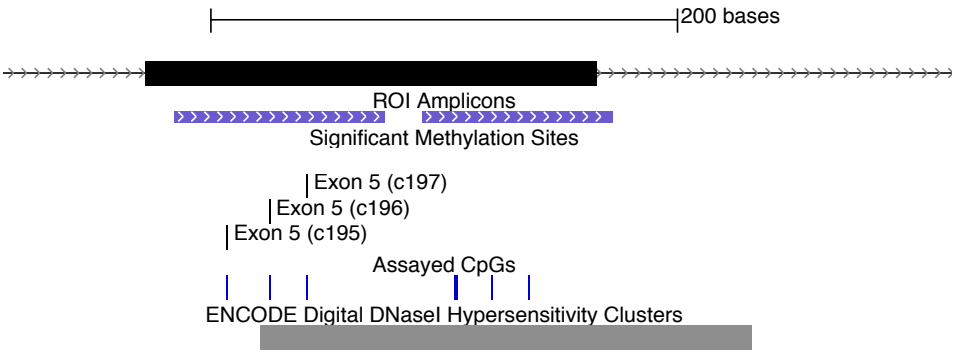


Figure 5.28 Exon 5 methylation sites in relation to DNase I hypersensitivity features from UCSC genome browser data [purple : ROI amplicons; black vertical bars: significant methylation sites; blue vertical bars: assayed CpG; ENCODE DNaseI Hypersensitivity Clusters track: signal indicated by grey and black boxes, where the darkness of the box is proportional to the maximum signal strength]

As shown in Figure 5.28, two of these methylation sites (c196 and c197) cover signal for DNase I hypersensitivity, as shown by the ENCODE DNase I hypersensitivity clusters track in the UCSC genome browser. However this data is based on a cell line (H1-hESC, embryonic stem cells) derived from tissue other than CNS, and only an indicator that this sequence may have some biological significance other than transcribing an exonic sequence.

5.6.8 CpG methylation correlation with Illumina Human Methylation 450k array

The CpG methylation data generated from the bisulphite sequencing was correlated with methylation data generated by colleagues on the same set of samples using the Illumina Human Methylation 450K array; to assess how the novel PCR based bisulphite sequencing approach presented here compared to commercial methylation array. Twenty-two CpG probes from the 450k array data cover the APP gene and 16 of these overlapped with the bisulphite sequencing data. Individual sample profile correlations of the 16 CpGs across the two data types are in good agreement with a Spearman's rho ranging from r=0.8-0.9 and p=<0.0001. As shown in Figure 5.43 the average methylation level at the 16 CpG sites is roughly the same across all 16 CpG sites, with the exception of CpG c12 c35 c58 c64 c116 c135 and c139 which show lower levels of methylation in the sequencing data and CpG c5 where the sequencing data show a increased level of methylation relative to the 450k array data.

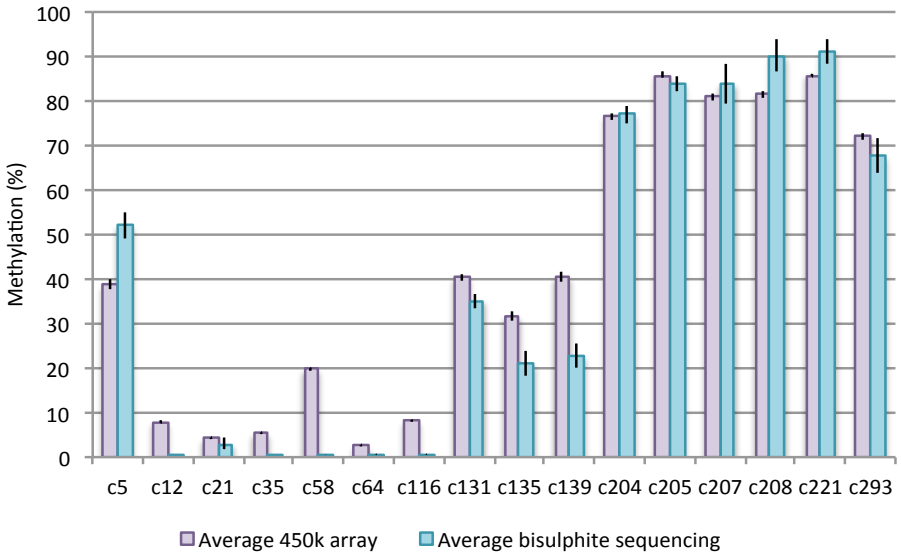


Figure 5.29 Bar chart of average sample CpG methylation levels by method [whiskers: 95% CI]

Variance in the sequencing data is much greater than that of the 450k array data as seen in Figure 5.29 and noted in Table 5.8 by the standard deviations. The higher variance in the sequencing data may partly explain why correlations of the individual CpGs across the two data types are weak and only significant for three of the CpGs tested, as noted in the last two

columns of Table 5.8. Also while both of the methods use bisulphite converted DNA, the array approach is based on probe hybridisation and the sequencing on PCR, thus they are both prone to different biases. Though there are differences in the data generated by the two methods, overall the sample profiles of methylation across the two methods are in good agreement and show that the data generated from the bisulphite sequencing approach presents comparable data to that generated by the commercial Illumina Human Methylation 450k array.

CpG	Feature location	450k Array		Sequencing		Correlation	
		Mean methylation (%)	Std. Dev.	Mean methylation (%)	Std. Dev.	rho	p
c5	5' region	38.61	5	52.1	13.29	0.169	0.122
c12	5' region	7.88	1.63	0.36	0.48	-0.098	0.403
c21	CGI	4.2	0.64	3.01	5	-0.01	0.932
c35	Exon1	5.36	0.79	0.43	0.28	-0.217	0.101
c58	CGI	19.7	2.42	0.22	0.12	0.214	0.052
c64	CGI	2.52	0.38	0.28	0.68	-0.108	0.353
c116	CGI	8.17	0.83	0.32	1.01	0.061	0.57
c131	IVS1 Enhancer 1	40.39	3.46	35.18	7.84	0.305	0.003
c135	IVS1 Enhancer 1	31.61	4.33	20.93	12.04	0.256	0.025
c139	IVS1 Enhancer 1	40.42	4.65	22.79	11.63	0.252	0.027
c204	Exon 7	76.47	3.27	76.94	9.87	0.169	0.108
c205	Exon 7	85.82	2.57	83.95	8.31	0.147	0.163
c207	Exon 7	80.86	3.32	83.77	20.06	0.021	0.848
c208	Exon 7	81.41	3.12	90.08	16.01	0.188	0.079
c221	Exon 9	85.8	1.41	91.18	9.89	0.078	0.596
c293	Exon 18	71.98	3.1	67.57	17.46	-0.018	0.872

Table 5.8 Average CpG methylation by method and Spearman's rho correlations

Associations of disease & CpG methylation across APP gene 450k array data

As with the bisulphite sequencing methylation calls, adjusted linear regressions (controlling for age sex and ApoE) were run for each of the array CpG features against AD status to test for any significant associations, which were then compared to the bisulphite sequencing results where applicable. Summarised β-coefficient & p-value are shown in Table 5.9, while the average levels of methylation at the 22 CpG sites by AD status are shown in Figures 5.30-5.31. From the linear regression results two array features showed significant associations of disease status with methylation levels. Firstly methylation at CGI AR3, a CpG site not covered by the bisulphite sequencing data, as it falls within the sequence of one of the ROI amplicons (immediate sequence 5' to the TSS) that dropped out. On average AD samples have 0.32% more methylation compared to controls (β = 0.323, SE = 0.146, p = 0.03). Secondly AD status is significantly associated with methylation at IVS1 AR10, a CpG site 3kb downstream of the CGI that was also not covered by the bisulphite sequencing data. On average AD samples have 2.1% more methylation compared to controls (β = 2.103, SE = 0.861, p = 0.017), however these p-

values do not withstand correction for multiple testing (Table 5.9). Also a direct comparison of the array and bisulphite sequencing could not be made for the two significantly associated sites from the array data, as they were not covered by the sequencing data.

Methylation CpG	Coefficient	Std. Err.	t	P-value	Corrected P-value	95% CI	
5' region CGI shore (AR11)	0.400	1.080	0.37	0.712	15.664	-1.751	2.552
5' region (c5)	2.173	1.385	1.57	0.121	2.662	-0.586	4.931
5' region (c12)	-0.369	0.450	-0.82	0.416	9.152	-1.266	0.529
CGI (c21)	0.130	0.178	0.73	0.466	10.252	-0.224	0.484
CGI (AR8)	0.332	0.217	1.53	0.130	2.860	-0.101	0.765
CGI (AR3)	0.323	0.146	2.21	0.030	0.660	0.031	0.614
CGI (AR2)	0.160	0.183	0.87	0.387	8.514	-0.206	0.525
Exon1 (c35)	-0.198	0.216	-0.92	0.362	7.964	-0.628	0.232
CGI (c58)	1.197	0.665	1.80	0.076	1.672	-0.127	2.522
CGI (c64)	-0.022	0.108	-0.21	0.836	18.392	-0.238	0.194
CGI (c116)	0.339	0.231	1.47	0.146	3.212	-0.121	0.799
IVS1 Enhancer 1 (c131)	-0.608	0.982	-0.62	0.538	11.836	-2.565	1.349
IVS1 Enhancer 1 (c135)	-0.837	1.210	-0.69	0.491	10.802	-3.249	1.574
IVS1 Enhancer 1 (c139)	-0.527	1.295	-0.41	0.685	15.07	-3.107	2.053
IVS1 (AR10)	2.103	0.861	2.44	0.017	0.374	0.388	3.818
Exon 7 (c204)	-0.653	0.913	-0.71	0.477	10.494	-2.473	1.168
Exon 7 (c205)	0.888	0.705	1.26	0.212	4.664	-0.517	2.294
Exon 7 (c207)	1.250	0.887	1.41	0.163	3.586	-0.516	3.017
Exon 7 (c208)	1.345	0.846	1.59	0.116	2.552	-0.341	3.03
Exon 9 (c221)	0.141	0.395	0.36	0.723	15.906	-0.647	0.928
IVS13 (AR22)	-0.188	0.691	-0.27	0.786	17.292	-1.566	1.189
Exon 18 (c293)	0.112	0.866	0.13	0.897	19.734	-1.613	1.837

Table 5.9 Summarised β-coefficient & p-value for test of disease marker associations with array CpG methylation [red p-value: <0.05; c methylation sites in brackets: CpG sites overlapping with bisulphite sequencing data; AR methylation sites in brackets: data unique to array and not covered by bisulphite sequencing data; corrected P-vales with a Bonferroni correction]

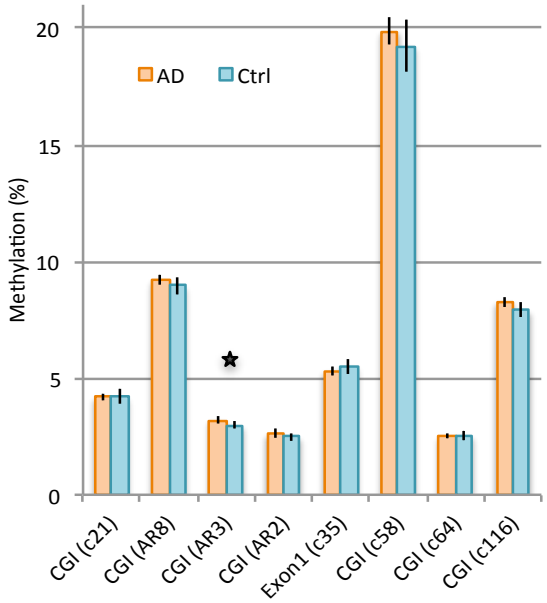


Figure 5.30 Bar chart of average array CpG methylation levels at the CGI by AD status [whiskers: 95% CI; black star: based on uncorrected p-value AD status is significantly associated with methylation at this site, see Table 5.9]

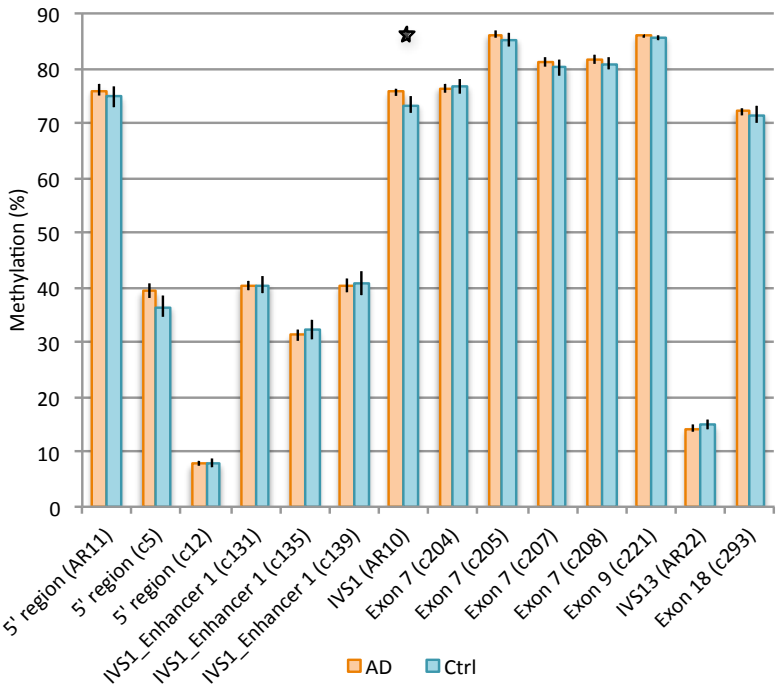


Figure 5.31 Bar chart of average array CpG methylation levels across the APP gene by AD status [whiskers: 95% CI; black star: based on uncorrected p-value AD status is significantly associated with methylation at this site, see Table 5.9]

5.6.9 SNP calls derived from the bisulphite sequencing data

Using bisulphite sequencing to generate base-pair resolution CpG methylation calls, also presents the opportunity to use the same sequencing data to call SNPs in the same samples. Initially SNP calling in the bisulphite sequencing data was attempted using a newly released programme, Bis-SNP (Liu et al. 2012), however it became apparent that the programme only calls SNPs reliably on bi-directional data, from one of the more standard bisulphite sequencing library protocols, such as the directional, MethylC-Seq (Lister et al. 2008) or the non-directional, BS-seq (Cokus et al. 2008) libraries. Whereas the bisulphite PCR and 100bp PE sequencing approach used here does not generate fully bi-directional PE data for the larger PCR amplicons (≥100bp), as it maximises on amplicon size (Figure 5.32b). Also only one of the original bisulphite converted strand (original-bottom [-] i.e. orientation of the APP gene) was amplified and sequenced (Figure 5.32a), thus excluding the ability to successfully call variants involving a cytosine in the minus [-] strand, as these correspond to either methylated CpGs or bisulphite converted cytosines in the original sequence and thus would result in ambiguous calls at that position.

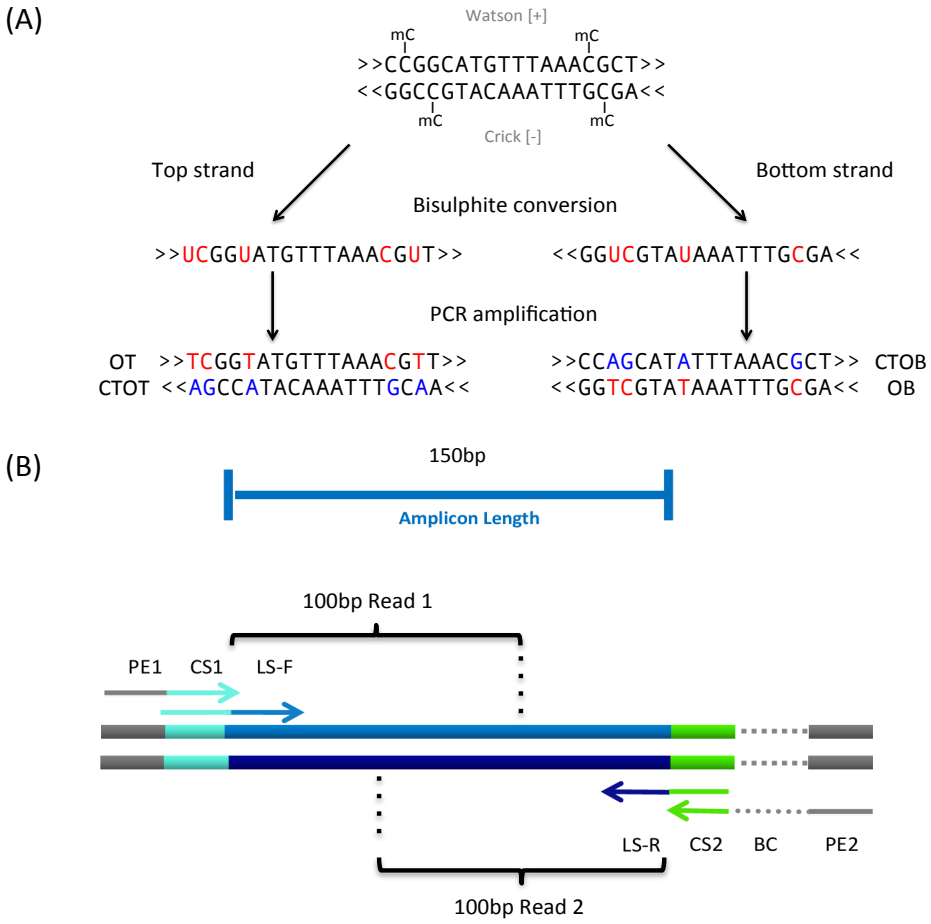


Figure 5.32 Illustration of bisulphite PCR & 100bp PE Illumina sequencing
[**A**: bisulphite conversion of genomic DNA gives rise to two uncomplimentary strands denoted as the top strand (+) and bottom strand (-), methylated cytosines are resistant to bisulphite conversion, whereas unmethylated cytosines are converted to uracil (U); subsequent PCR amplification can produce up to four potentially different DNA fragments for any given locus; mC: 5-methylcytosine; OT: original top strand; CTOT: strand complementary to the original top strand; OB: original bottom strand; CTOB: strand complementary to the original bottom strand; adapted from Krueger et al. 2012; **B**: schematic of sequence covered by paired-end reads (as described in chapter 2 section 2.18); PE: paired end Illumina sequencing adaptor sequence; CS: common sequence between primary and secondary PCR primers; LS: locus specific primer sequence; BC: barcode sequence; average PCR amplicon size is 150bp and paired-end reads 1 and 2 sequence 100bp from either end of the original bottom bisulphite converted strand, i.e. in the orientation of the APP gene]

Following the unsuccessful SNP calling attempt with the Bis-SNP package, SNPs were successfully called using the Genome Analysis Tool Kit (GATK) package, with the proviso that only calls at bases not involving a guanine (in the [+] orientation) were likely to be unambiguous. From the successfully called GATK SNPs, a filtered list of known SNPs with dbSNP build 135 rs numbers resulted in 20 SNP calls. The 20 SNPs are listed in Table 5.10, along with their respective rs numbers and the bisulphite sequencing ROI amplicon on which they reside.

SNP ID	Bisulphite sequencing ROI amplicon	rs number	Ref/Alt alleles	Comparable SNP array data
bSNP 1	IVS4 (f85)	rs6516719	C/T	+
bSNP 2	IVS11 (f107)	rs440666	T/C	+
bSNP 3	IVS1_Enhancer 3 (f61)	rs2830071	C/T	+
bSNP 4	IVS1_Enhancer 3 (f62)	rs2830070	T/C	+
bSNP 5	IVS2_Open Chromatn 3 (f77)	rs2830052	C/T	+
bSNP 6	IVS13 (f6d)	rs2829996	T/A	+
bSNP 7	IVS17_Open Chromatin 10 (f27)	rs214485	C/A	+
bSNP 8	IVS13 (f6)	rs190392435	T/C	-
bSNP 9	IVS1_Open Chromatn 2 (f65)	rs181441239	T/C	-
bSNP 10	IVS13 (f6d)	rs1701001	C/T	+
bSNP 11	Exon 9 (f99)	rs148257054	T/C	-
bSNP 12	IVS1_Enhancer 1 (f52)	rs147054581	C/T	+
bSNP 13	IVS16 (f20)	rs144506349	T/C	-
bSNP 14	IVS1 (f46)	rs139326696	T/A	-
bSNP 15	IVS7_Open Chromatn 6 (f94)	rs139128431	T/C	+
bSNP 16	IVS3_Open Chromatn 4 (f81)	rs12626415	C/T	+
bSNP 17	IVS1_Enhancer 2 (f57)	rs117982323	C/A	+
bSNP 18	IVS10 (f100)	rs115962715	T/C	-
bSNP 19	Exon 14 (f16)	rs112263157	T/C	-
bSNP 20	IVS11_Open Chromatin 7 (f105)	rs112248226	T/C	-

Table 5.10 List of SNPs genotyped from bisulphite sequencing data
[b prefix to SNP numbers denotes the SNP call is derived from bisulphite sequencing data]

Correlation of SNP calls with known SNP array data

To check the validity of the SNP calls derived from the bisulphite sequencing data, the generated genotypes were cross-referenced with known genotypes from SNP array data. SNP array data was available for 55 of the samples that were genotyped above and 12 of the SNPs genotyped by bisulphite sequencing, as identified in Table 5.10 overlapped with the array SNPs. Results of the comparison showed that 97.3% of the bisulphite sequencing derived genotypes matched the known genotypes from the SNP array data.

methQTL analysis

With the confidence that the SNP genotypes from the bisulphite sequencing data were reliable, sample genotypes for these 20 SNPs in Table 5.10, along with another 43 SNP genotypes from the SNP array data that were identified to be ± 100bp from any given methylation site were then tested for possible associations with the CpG methylation data. For this linear regressions were carried out on the methylation level (%) as the outcome variable and SNP genotype values: 0 = homozygous for the reference allele; 1 = heterozygous; and 2 = heterozygous for the alternative allele (along with covariates age and sex) as the predictor variable. In total 63 SNPs were tested against 271 methylation sequences. Twenty-one associations of SNP genotypes

with methylation levels were identified at a significance level of p = 0.05 and are sorted and listed by p-value in Table 5.11.

Methylation site tested for methQTL	SNP ID	rs number	Ref/Alt alleles	P-value	Corrected P-value
IVS4 (c194)	bSNP 1	rs6516719	C/T	5.76×10^{-30}	1.56×10^{-27}
IVS12 (c245)	aSNP 19	rs417676	G/A	2.00×10^{-16}	5.42×10^{-14}
IVS12 (f4)	aSNP 19	rs417676	G/A	6.74×10^{-16}	1.83×10^{-13}
IVS4 (f85)	bSNP 1	rs6516719	C/T	3.74×10^{-14}	1.01×10^{-11}
IVS2_Open Chromatin 3 (f77)	bSNP 5	rs2830052	C/T	7.70×10^{-09}	2.09×10^{-06}
IVS17_Open Chromatin10 (c289)	bSNP 7	rs214485	C/A	5.22×10^{-06}	1.41×10^{-03}
5' region (f93)	aSNP 7	rs139885956	G/A	8.48×10^{-04}	2.30×10^{-01}
5' region (c6)	aSNP 7	rs139885956	G/A	3.30×10^{-03}	8.94×10^{-01}
5' region (c7)	aSNP 7	rs139885956	G/A	4.82×10^{-03}	1.31
CGI (c80)	aSNP 53	rs117107800	C/G	7.48×10^{-03}	2.03
IVS1_Enhancer 4 (c159)	aSNP 48	rs13048354	C/T	7.69×10^{-03}	2.08
CGI (c71)	aSNP 53	rs117107800	C/G	7.71×10^{-03}	2.09
CGI (c75)	aSNP 53	rs117107800	C/G	2.13×10^{-02}	5.77
CGI (c100)	aSNP 28	rs469420	A/G	2.16×10^{-02}	5.85
IVS13_Open Chromatin 8 (c259)	aSNP 1	rs116967391	G/C	2.31×10^{-02}	6.26
CGI (c79)	aSNP 53	rs117107800	C/G	2.68×10^{-02}	7.26
IVS1_Open Chromatin 2 (c154)	bSNP 9	rs181441239	T/C	2.86×10^{-02}	7.75
CGI (c109)	aSNP 21	rs45473297	G/A	3.29×10^{-02}	8.91
CGI (c103)	aSNP 21	rs45473297	G/A	4.46×10^{-02}	12.09
IVS13 (f6)	bSNP 6	rs2829996	T/A	4.59×10^{-02}	12.44
IVS1_Alu (f66)	aSNP 48	rs13048354	C/T	4.68×10^{-02}	12.68

Table 5.11 Summarised list of significant p-values from methQTL analysis
[a: prefix to SNP numbers denotes the SNP call is derived from SNP array data; b: prefix to SNP numbers denotes the SNP call is derived from bisulphite sequencing data; red p-value:<0.05; corrected P-vales with a Bonferroni correction]

The quantile-quantile plot of the observed p-values against the expected (Figure 5.33) shows that most of the p-values are the same as the expected and lie on top of the null distribution, implying that there are no biases in the data and a low possibility of false positive associations. Whereas the top few p-values deviate significantly from the null distribution. Following a Bonferroni correction of these p-values to account for multiple testing, the top six associations remain significant at a $p \leq 0.05$ level (Table 5.11).

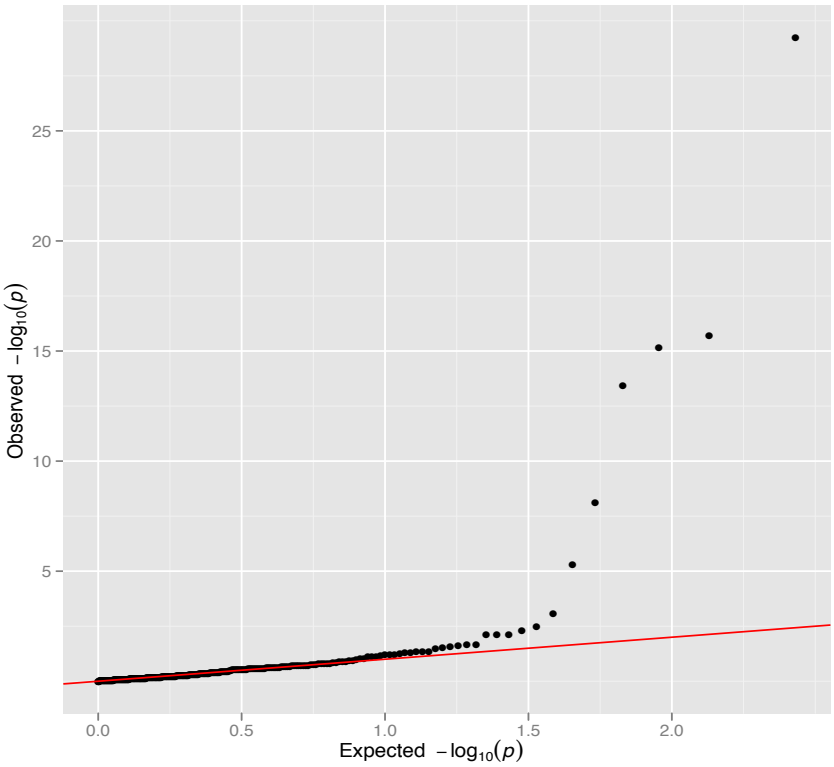


Figure 5.33 quantile–quantile plot of the 271 test of association p-values
[black dot: represents an observed statistic [defined as the −log10(P)] versus the corresponding expected statistic; red line : corresponds to the null distribution]

The finding of the association of bSNP1 with both IVS4 CpG c194 and the ROI sequence f85 in which this CpG resides is essentially the same association, where the association with the ROI sequence f85, is likely to be driven by the CpG c194 at this ROI. A similar result is seen where the association of aSNP 19 occurs with both the IVS4 ROI sequence f4 and the CpG c245 within this ROI sequence and the association with the ROI sequence f4, is probably driven by the CpG c245 that resides at the f4 ROI.

Association of bSNP 1 (rs6516719) with CpG methylation at IVS4 c194

The location of bSNP 1 (rs6516719, C/T) in relation to CpG c194 shows that the SNP it self is actually at the cytosine position of the CpG site, as shown on the hg18 forward strand of the APP gene in Figure 5.34 and that the distribution of c194 methylation by bSNP 1 genotype shows a clear demarcation of the methylation levels by genotype, as would be expected for a CpG SNP (Figure 5.35). With the C/C genotype showing high levels of methylation close to 100%, while the T/T genotype shows low levels of methylation down to 0% and the heterozygote genotype, C/T with methylation levels falling in between the two heterozygous geneotypes.

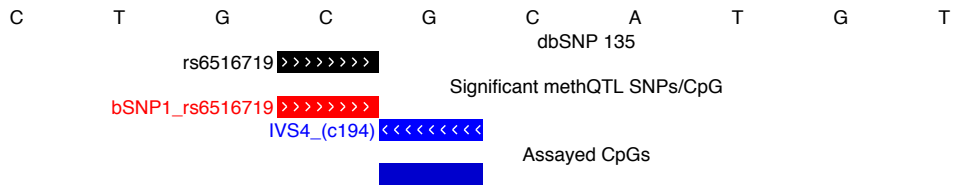


Figure 5.34 Schematic of bSNP 1 and IVS4 CpG c194
[dbSNP 135 track: showing location of SNPs (in black) at the given sequence; significant methQTL SNPs/CpGs track: shows SNP (in red) associated with CpG (in blue); DNA sequence is orientated to the forward strand [+] relating to the SNP]

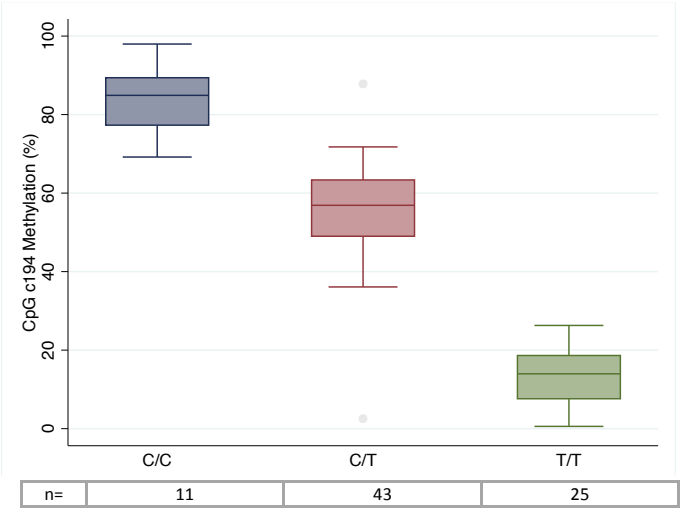


Figure 5.35 Boxplot of IVS4 CpG c194 methylation by bSNP 1 genotype
[whiskers: 95% CI; n per group tabulated at bottom of figure]

Association of bSNP 5 (rs2830052) with CpG methylation at IVS2 open chromatin 3 ROI CpG f77

The location of bSNP 5 (rs2830052, C/T) in relation to CpG f77 shows that the SNP is 62 bp from the assayed single CpG site in ROI f77 (Figure 5.36). Thought the CpG site it self is actually at a position where a SNP (rs73340764, G/A) resides and there was no genotype data for this rs73340764 SNP in the SNP array dataset or that derived from the bisulphite sequencing to rule out the possibility of this SNP occurring at this CpG in our dataset, this SNP is rare and is reported to have a minor allele frequency (MAF) of 0.5% in the 1000 genome phase 1 population as noted on the dbSNP database.

Though the possibility that this SNP (rs73340764) may influence methylation at the ROI CpG f77 cannot be completely ruled out, the current data suggest that association of bSNP 5 with methylation at CpG f77 signifies an example of cis-methQTL. The relationship of bSNP 5 genotype and methylation is shown in Figure 5.37, where the C/C genotype confers higher methylation levels (mean ~90%) at CpG f77 whereas the T/T genotype on average presents with

lower methylation (mean ~55%) and the heterozygous C/T falls in between these values with average methylation ~70%. To test the relationship of bSNP 5 and methylation at CpG f77 for allele-specific methylation, the methylation percentage values were correlated with their corresponding percentage values of allele ratios as calculated from the SNP call read depth, which indicates a significant correlation (rho = 0.54, p = <0.001). As shown in Figure 5.36, this region lies within the sequence identified as IVS2 open chromatin 3 and though there was no significant association of disease status and methylation in this region there is signal for TF occupancy by a number of different TFs, as indicated by the ENCODE transcription factor ChIP-seq track. However the overlapping TF (c-Jun and BAF155) occupancy data are based on cell lines derived from tissue other than CNS and therefore not directly relevant and only indicative of a general role in gene regulation.

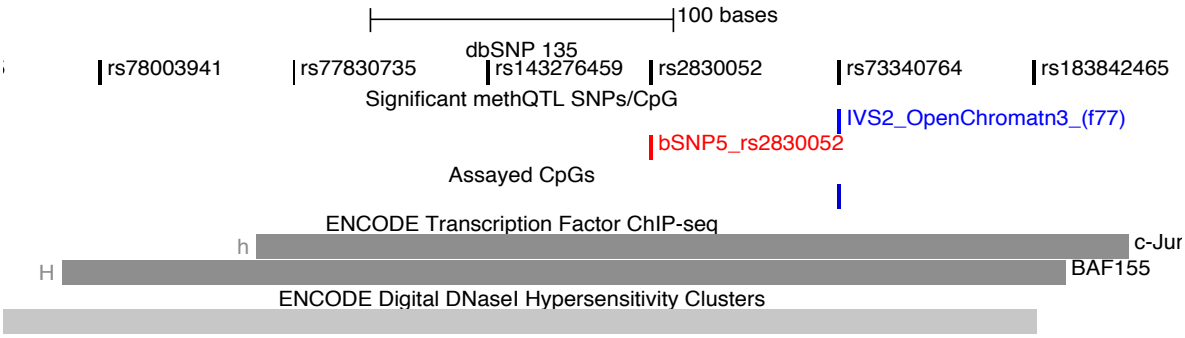


Figure 5.36 Schematic of bSNP 5 and IVS2 open chromatin 3 ROI CpG f77
[dbSNP 135 track: showing location of SNPs (in black) at the given sequence; significant methQTL SNPs/CpGs track: shows SNP (in red) associated with CpG (in blue); small blue vertical bar: assayed CpG; ENCODE TF ChIP-seq track, TF occupancy signal indicated by grey and black boxes, where the darkness of the box is proportional to the maximum signal strength, the TF name is shown to the right of the box, while the letters to the left of the box represent the cell lines where a signal is detected, see abbreviations section for full cell line names; ENCODE DNaseI Hypersensitivity Clusters track, signal indicated by grey and black boxes, where the darkness of the box is proportional to the maximum signal strength]

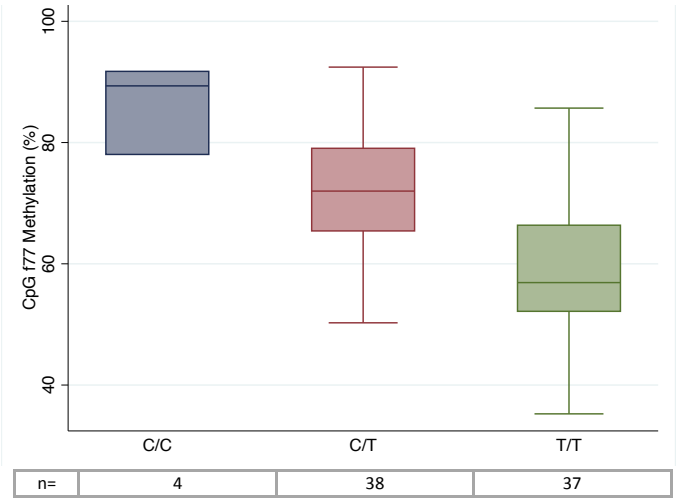


Figure 5.37 Boxplot of IVS2 ROI CpG f77 methylation by bSNP 5 genotype
[whiskers: 95% CI; n per group tabulated at bottom of figure]

Association of bSNP 7 (rs214485) with CpG methylation at IVS17 open chromatin 10 CpG c289

The location of bSNP 7 (rs214485, C/A) in relation to CpG c289 shows that the SNP it self is actually at the cytosine position of the CpG site, as shown on the hg18 forward strand of the APP gene in Figure 5.38a. Within our dataset only the heterozygous C/A and homozygous A/A genotypes are present. Methylation levels at CpG c289 for the C/A individuals range from 0.23-6.82%, while A/A homozygotes present levels ranging 0-0.27% (Figure 5.39). Also bSNP 7 and CpG c289 are within an open chromatin sequence that may hold some regulatory potential (Figure 5.38b).

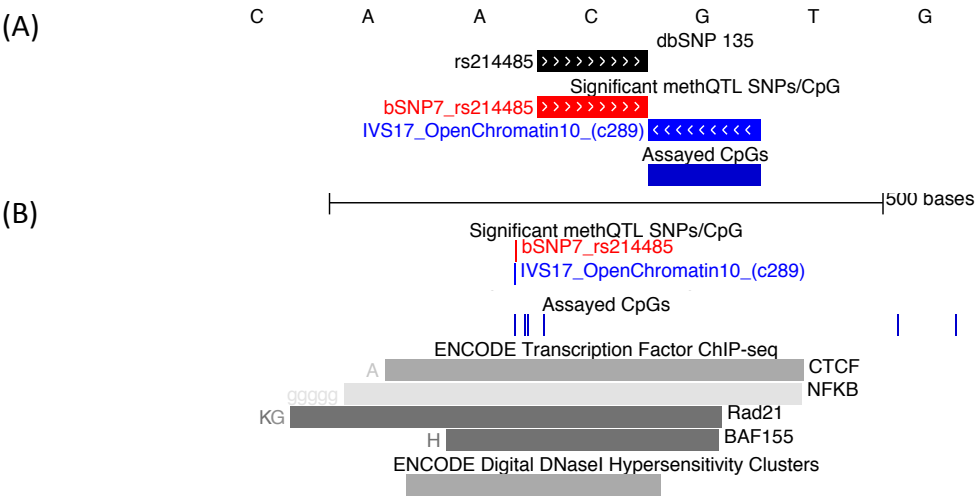


Figure 5.38 Schematic of bSNP 7 and IVS17 open chromatin 10 CpG c289
[**A**: dbSNP 135 track: showing location of SNPs (in black) at the given sequence; significant methQTL SNPs/CpGs track: shows SNP (in red) associated with CpG (in blue); DNA sequence is orientated to the forward strand [+] relating to the SNP reference allele; **B**: schematic of IVS17 open chromatin region; ENCODE TF ChIP-seq track, TF occupancy signal indicated by grey and black boxes, where the darkness of the box is proportional to the maximum signal strength, the TF name is shown to the right of the box, while the letters to the left of the box represent the cell lines where a signal is detected, see abbreviations section for full cell line names; ENCODE DNaseI Hypersensitivity Clusters track, signal indicated by grey and black boxes, where the darkness of the box is proportional to the maximum signal strength]

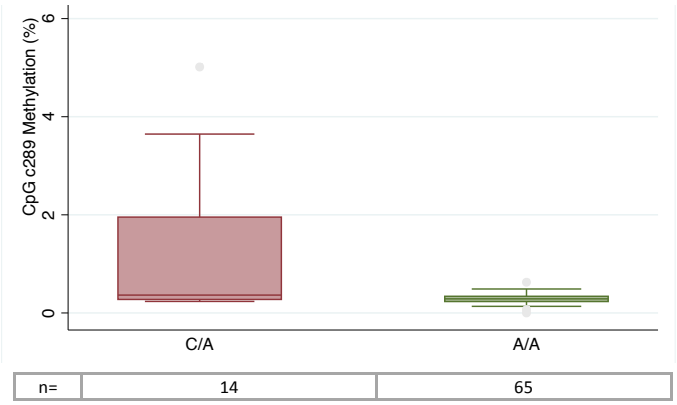


Figure 5.39 Boxplot of IVS17 CpG c289 methylation by bSNP 7 genotype
[whiskers: 95% CI; n per group tabulated at bottom of figure]

Association of aSNP 19 (rs417676) with CpG methylation at IVS12 CpG c245

From the SNP array data, aSNP 19 was associated with methylation at IVS12 CpG c245, where the bSNP 19 (rs417676, G/A) is at the guanine position of the CpG c245 as shown on the hg18 forward strand of the APP gene in Figure 5.40. Only the heterozygous G/A and homozygous A/A genotypes are present, with the G/A genotype showing levels of methylation ranging from ~20-70%, while the homozygote A/A genotype exhibits 0% methylation levels. Also this CpG showed a spike as a potential differentially methylated site (Figures 5.22-5.24), which is likely to be the result of the lower levels of methylation in the G/A heterozygous AD samples and the higher number of AD samples with the A/A genotype relative to the controls (Figure 5.41), though when AD status was added and tested as a further predictor variable this association proved to be negative (β = -4.631, SE = 2.372, p = 0.057).

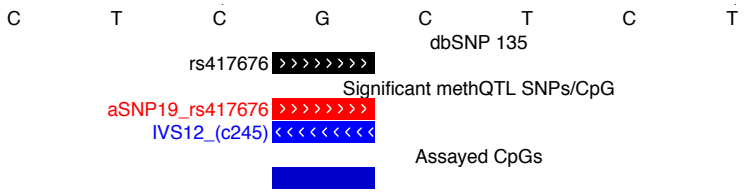


Figure 5.40 Schematic of aSNP 19 and IVS12 CpG c245
[dbSNP 135 track: showing location of SNPs (in black) at the given sequence; significant methQTL SNPs/CpGs track: shows SNP (in red) associated with CpG (in blue); DNA sequence is orientated to the forward strand [+] relating to the SNP reference allele]

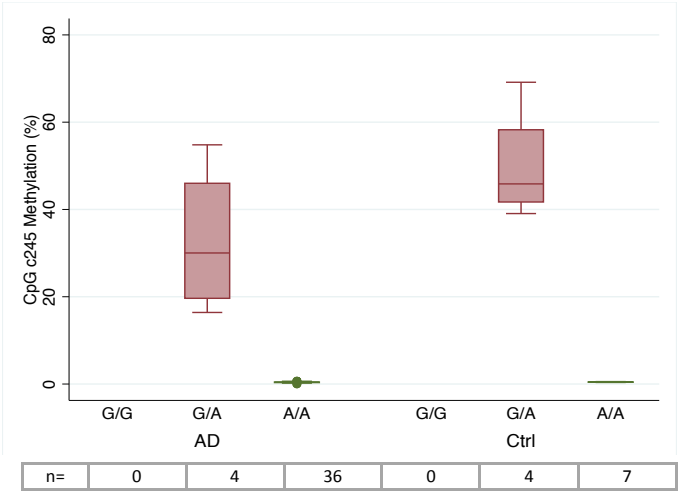


Figure 5.41 Boxplot of IVS12 CpG c245 methylation by aSNP 19 genotype and AD status
[whiskers: 95% CI; n per group tabulated at bottom of figure]

Given the limitation that no SNPs could be called at guanine positions on the bisulphite sequencing (on the hg18 + strand) and that the SNP array data only covered SNPs up to dbSNP build 129, all the CpG sites identified as having a significant association of disease status (c1, c195, c196 and c197) were checked for possible G/A SNPs, up to and including dbSNP build 135. However no such SNPs were identified at these CpG sites.

5.7 Discussion

APP mutations and the neuropathology of individuals with three copies of this gene show that an alteration in either the amino acid sequence of Aβ or the amount of Aβ produced is sufficient to cause AD and supports the notion that increased expression of APP can cause an increase in concentration of Aβ40/42 peptides and subsequent Aβ deposition that is critical for AD pathogenesis. While APP expression produces a heterogeneous product in the form of three main isoforms in the brain, APP gene regulation and transcription is complex involving both proximal and distal regulatory sequences that may be influenced by DNA methylation.

Methylation status of the APP gene was assessed by bisulphite sequencing. The data presented in this chapter demonstrate the use of Fluidigm access array 48.48 microfluidics in generating a barcoded bisulphite PCR sequencing library covering 22kb of the APP gene region. While the Fluidigm access array system has been widely used for non-bisulphite parallel sequencing (Hollants et al. 2012), it has not previously been described for bisulphite sequencing and in doing so the data presented here show that it can be applied to the simultaneous testing of 96 samples using barcoded adaptors to generate high depth base-pair resolution methylation data. Furthermore cross platform comparison of the APP bisulphite sequencing methylation data with data from the Illumina Human Methylation 450K array showed that the methylation profiles of overlapping CpG sites was in good agreement, as has been reported elsewhere (Bibikova et al. 2011).

Average methylation levels across the entire set of CpGs tested show AD samples to have 32.6% methylation and control samples to have 34.8% methylation at the APP gene region and though this difference is not statistically significant it falls in line with the observation of decreased global 5mC measured in AD brain samples compared to controls in chapter 3. However an attempt to further investigate the trends seen for global methylation changes at consensus Alu sequences in chapter 3, specifically at the APP locus was unsuccessful due to the challenges of amplifying repeat sequences.

While there are significant sex differences in average methylation levels at five separate sequences (out of 114 sequences tested) across the APP gene from IVS1, IVS2, IVS13 through to exon 18, with males having ~5% less methylation than females; no such differences have been previously reported for APP gene methylation. Whereas, Wang et al. (2008) observed there to be no significant sex differences in APP promoter methylation in either brain samples or lymphocytes. Similarly there was no significant sex difference detected in any of the promoter/CGI sequences test in the current data.

There are marginally significant associations of average methylation levels with age, at seven separate sequences and while two sequences in IVS3 indicate hypomethylation with age, with a 0.15% decrease per year; the other five sequences (IVS3, exon 5, exon 11, IVS11 and exon 17) indicate hypermethylation with age averaging around a 0.26% increase in methylation per year. These results relating to age and methylation, are not from the same individuals across different time points and thus only an indication of a relationship of APP methylation with time. Also there was no significant association of age with methylation within the promoter/CGI region of APP and while this is at odds with the age-dependent methylation changes previously reported for the APP gene promoter (Tohgi et al. 1999a ; Nagane et al. 2000), the promoter/CGI sequence covered in those studies falls within one of the ROI amplicons that failed to amplify successfully. Also no other studies measuring APP promoter methylation to date, have covered this same sequence, possibly owing to technical difficulties with this stretch of sequence as described by Nagane et al. (2000).

Overall average methylation levels at different features (CGI, exons, introns, 5'region, 3' UTR, enhancer and open chromatin sequences) within the APP gene are similar for both cases and controls and while the APP CGI feature shows the canonical pattern of a fully hypomethylated promoter. Before correction for multiple testing, AD status was observed to be significantly associated with methylation levels at one CpG site (c1) within the 5' region CGI shore. This association was in a positive direction indicating hypermethylation with disease. Consistent with other studies of DNA methylation in the proximal promoter region of APP (Wang et al. 2008; Barrachina & Ferrer 2009), there was no observation of differential methylation between cases and controls within the sequence range -515/-257 bp upstream of the TSS.

Two previous studies have investigated CpG methylation within the 5' region, one reported there to be no methylation differences in the frontal cortex and hippocampus of cases and controls (Barrachina & Ferrer 2009), covering -2572/-2108 bp, while the other identified a differentially methylation CpG site (-1184 bp) at various brain regions in non-demented

individuals (Rogaev et al. 1994) and though the same exact sequences were not examined in the current study, there was observation of methylation difference at one CpG site (-2781 bp, 5' CGI shore c1) neighbouring these previously tested sequences. These results cannot confirm the finding of Barrachina & Ferrer (2009), however they partially corroborate the finding by Rogaev et al. (1994). This observation of a 5' region DMR may relate to tissue-specific methylation (Doi et al. 2009) in brain (Davies et al. 2012) and possibly expression as has been shown to be important for other types of tissue (Irizarry et al. 2009).

Before correction for multiple testing, significant differentially methylated CpGs were also observed at exon 5. Where AD brain on average exhibits ~5% more methylation at exon 5 compared to controls and though the relationship between gene body methylation and transcription is obscure, a growing body of evidence suggests that methylation outside of the core CGI is also likely to be involved in transcriptional regulation (Brenet et al. 2011). For example, Maunakea et al. (2010) have shown methylated intragenic regions to correlate with higher levels of gene transcription in human brain, which is also documented in cell transfection experiments (Bauer et al. 2010). While intragenic methylation may lead to the modification of the transcriptional efficiency of RNA polymerase II and the transcriptional complex (Lorincz et al, 2004), more recently intragenic DNA methylation was demonstrated to prevent CTCF binding, which facilitates RNA polymerase II elongation and promotes exon skipping (Shukla et al. 2011). However there are no co-occurring separate associations of total APP protein levels at these four sites.

A small number of known SNPs genotyped using the bisulphite sequencing data along with previously generated array SNP data were tested for associations with CpG methylation. However only four significant potential methQTLs were identified and of these three sites revealed to be CpG SNPs, with only one site showing a cis-methQTL effect. This cis-methQTL of SNP rs2830052 with CpG methylation within intron 12 open chromatin sequence also exhibits a significant correlation of between the allele ratios and methylation levels indicative of allele-specific methylation. Though the data for the methQTLs analysis were limited by the small number of known SNPs successfully called within the bisulphite sequencing data, owing to the fact that only one of the original bisulphite converted strands was sequenced; methQTL analysis as performed here highlights the need to account for CpG SNPs in DNA methylation studies.

There are a number of limitations that need to be considered in drawing conclusions from this study. Firstly, the p-values for significant associations of methylation with disease do not remain significant after correction for multiple testing and so should be interpreted with caution and

need to be replicated in another sample; and while the 5' region methylation difference is at single CpG site, the exon 5 potential DMR spans over a sequence with three associated CpGs that may represent a more robust association of disease status with methylation levels. Ideally a larger number of samples with more controls will be required to achieve a significance level withstanding correction for multiple testing.

Secondly a major inherent limitation of this study is that measurements were performed on DNA extracted from brain homogenates that included both normal and damaged neurons and prone to varying populations of glial and other cells accountable to the disease process. One option to over come this would be to use cell populations selectively collected using a method such as Fluorescence-Activated Cell Sorting (FACS) or Laser Capture Microdissection (LCM). Furthermore total APP protein levels previously measured by Western blotting only provide a crude measure of the end product following post-transcriptional and translational processing and unlikely to reflect transcriptional output alone. Again better characterised cell populations with more accurate mRNA measurements would be desirable to detect methylation-associated changes in APP expression.

Lastly a limitation of the current dataset is that it only measures total DNA methylation (5mC+5hmC) levels and does not distinguish modified cytosines as being either 5mC or 5hmC. Given the observed differences in the global levels of these two modification at the STG brain area in AD (in chapter 3), along with their proposed distinct roles; separate quantification of these modification at the APP gene locus may further delineate DNA methylation differences related to disease. While sample preparation methods allowing the separation of these two DNA methylation marks for downstream sequencing were not available at the time of planning this study, DNA conversion methods specific to 5hmC are now available, which could easily be applied to the sequencing approach used in the current study (Booth et al. 2012; Yu et al. 2012).

DISCUSSION

The primary aim of this thesis was to examine case-control differences in DNA methylation using post-mortem tissue covering brain areas affected by AD pathology and those that are spared of AD pathology or associated with behavioural phenotypes. Three studies were presented in this thesis that explored both global DNA methylation levels and DNA methylation at particular sequences of two candidate genes.

6.1 Summary of findings

Chapter 3 explored global genome-wide DNA methylation levels using a number of different measures and techniques to ascertain if there are measureable methylation differences in AD pathology affected brain areas and non-AD pathology affected areas across cases and controls. Simultaneous measurement of both 5mC and 5hmC methylation was successfully achieved with a modified version of the LUMA assay (GLUMA). Hypomethylation of 5mC and hypermethylation of 5hmC were observed in AD brain relative to controls, across both AD pathology affected and unaffected brain areas. Also the observed inverse relationship in the levels of 5hmC and 5mC is consistent with the notion that 5hmC serves as an intermediary form for demethylation of 5mC. While age was separately associated with a decrease in LINE1 methylation and an increase in 5hmC methylation levels.

Chapter 4 examined the association between the serotonin transporter (SERT) gene promoter DNA methylation and mRNA abundance across seven different brain areas in relation to AD status and additionally assessed the associations of 5HTTLPR genotype, the presence of stressful life events (SLE) and depression in a subset of these AD samples. While AD status was tentatively associated with the average methylation levels of the SERT CGI in brain area BA10 (a brain area associated with behaviour), methylation at a specific sequence within the CGI showed a similar relationship with Braak staging. Although these associations of disease with methylation levels did not extend to SERT mRNA levels. The observed methylation levels at the SERT CGI are extremely low and below the technical detection level of the Sequenom MassARRAY EpiTYPER platform and any measurable differences are very small and not likely to be significant, the analysis in this chapter was purely exploratory.

Further analysis on the subset of AD patients with additional clinical and behavioural measures showed AD patients with the 5HTTLPR S/S genotype to having on average lower mRNA expression levels than those with either of L/L, L/S genotypes at the frontal cortex areas. Also these preliminary exploratory findings demonstrated an association of L/L, L/S genotypes with depression. Interestingly, while there was no effect of 5HTTLPR genotype on any measure of CGI DNA methylation, both presence of SLE and higher depression scores are associated with higher levels of CpG methylation within the 3' portion of the CGI.

Chapter 5 examined the association between APP gene region DNA methylation and AD status along with APP protein levels in an AD pathology affected brain area (STG). Bisulphite sequencing data was successfully generated using a bisulphite PCR based parallel sequencing methodology. Whilst age was associated with average methylation levels at seven separate intragenic sequences no such association was found to occur within the promoter/CGI region of APP. Overall average methylation levels at different feature sequences (CGI, exons, introns, 5'region, 3' UTR, enhancer and open chromatin sequences) within the APP gene are comparable for cases and controls. AD status is associated with methylation status of several CpGs. Highlighting differentially methylated sites within the 5' region CGI shore and exon 5.

SNP calls from the same bisulphite sequencing data were also explored for association with nearby CpG methylation levels and identified four methQTLs, three of which revealed to be CpG SNPs, while one site of cis-methQTL effect showed allele-specific methylation.

Essentially the DNA methylation studies presented in this thesis loosely represent different aspects of AD risk, which include aging (explored through global DNA methylation), depression (and SERT promoter methylation) and amyloid-β deposition in brain (by exploring APP gene region methylation). Any of the observed DNA methylation changes on their own or in combination could possibly alter the susceptibility of aged brain in developing AD. The measures of global DNA methylation presented in Chapter 3 indicate changes with age; specifically elevated levels of 5hmC in AD brain may embody accelerated changes with age that also relate to other environmental risk factors of AD such as increased hyperhomocysteinemia and dietary deficiencies in folate or B vitamins. While the studies cover slightly disparate themes, the findings from the global methylation study could conceivably relate to APP gene-body methylation differences given the recent reports suggesting a role for 5hmC in neurodegenerative process (Song et al. 2010), along with intragenic enrichment leading to increased transcription (Ficz et al. 2011; Khare et al. 2012) and lower levels at alternatively spliced transcripts (Khare et al. 2012) and though 5hmC content was not specifically measured

at the APP locus one might speculate it forms a greater portion of the total gene-body methylation in AD brain relative to controls based on differences observed at a global level.

6.2 Strengths & limitations

A strength of these studies is that methods used in each of the studies have been both sensitive and current. While the use of Sequenom's MassARRAY MALDI-TOF mass spectrometry platform allowed accurate measurements of CpG methylation of at the SERT gene CGI in Chapter 4; highly parallel bisulphite sequencing was assessed and employed to detect base-pair resolution CpG methylation across the APP gene region in Chapter 5. Furthermore the GLUMA method was developed in Chapter 3, which provided an accurate means by which the global content of both 5mC and 5hmC could be assessed in a single assay. Both the GLUMA technique along with the Alu and LINE1 global methylation assays used the pyrosequencing platform that provides highly sensitive quantitative measurements.

The numbers of samples tested throughout the thesis have been reasonable, give the scope of these studies and analysis has incorporated all relevant clinical and demographic data where possible. The SERT gene and depression study (Chapter 4) utilised additional cohort data to include a large number of in-life measures including treatment with medication, which could possibly affect DNA methylation analysis. However this study of a small subset of AD patients with depression was limited in its power to detect certain established associations and any interactions between predictor variables.

As already mentioned this work is not without its limitations; the remainder of this section will focus on those limitations that are general to all the studies presented (Chapters 3 to 5) in this thesis and consider possible approaches to future studies in light of these methodological limitations.

At best the changes in DNA methylation levels observed here and in the literature are as small as 5% and even lower when examining hypomethylated promoters for possible DMRs. These small differences are liable to the requirement of a large number of samples to reach robust significance, thus one of the limitations of this thesis is that some of the analyses carried out are possibly underpowered and sampling bias cannot be ruled out where a limited number of samples were tested. Further to this the p-values for significant associations of methylation with disease have either, not been corrected for multiple testing or do not withstand such

correction, so should be taken as preliminary and need to be replicated in another sample possibly using a different technique. Ideally a larger number of samples with more controls (as control post-mortem samples seem to be a major limitation in these studies) will be required to achieve a significance level withstanding correction for multiple testing.

Another major inherent limitation of these studies concerns the use of post-mortem brain homogenates in that they represent a collection of different cell types including neurons and glia cells, which may vary from sample to sample at two levels. Firstly whilst care is taken in matching the same specific brain segment between different brain samples when excising a tissue segment of interest from the whole post-mortem brain, varying portions of grey and white matter can be introduced at the sample collection stage. Secondly diseased brain shows wide spread neuronal atrophy further compounding the differences between patient samples. Thus DNA extracted from case-control brain homogenates is likely to represent varying cell compositions and given that DNA methylation shows tissue-specific profiles, it is not know for certain whether the observed DNA methylation differences between cases and controls represent true disease related differences at the neuronal level or if they are simply a consequence of the neurodegenerative processes associated with AD, that result in differential cell death of neurons and gliosis, or possibly influenced by tissue sampling variation.

Ultimately identified disease associated DNA methylation changes at specific gene loci need to be functionally assessed in relation to gene expression levels in the same tissue to hold any biological relevance to the disease phenotype. While SERT mRNA levels (Chapter 4) and APP protein level (Chapter 5) were assessed for associations with DNA methylation these specific measures had limitations of their own, as already covered in their respective Chapter discussions. Further to this, these specific measures may not accurately reflect gene expression as they are also derived from post-mortem brain homogenates likely to represent a mixture of transcript variants and isoforms that are specific to a particular cell type. Again better characterised cell populations with more accurate mRNA measurements are likely to be required to detect methylation-associated changes in expression. For these reasons methods that enable the study of specific cell populations are likely to be key in deciphering the role of DNA methylation and other epigenetic marks in AD.

Possible solutions include the use of cell populations selectively collected using a method such as Laser Capture Microdissection (LCM) or Fluorescence-Activated Cell Sorting (FACS), as used by Siegmund et al. (2007) to study neuron specific methylation differences in AD post-mortem brain. While the cells collected using these methods provide only a small number of cells and

thus low yields from DNA extraction which may not be suitable to the requirements of the methods used in this thesis, protocols such as, Reduced Representation Bisulfite Sequencing (RRBS) have been optimized to used as little as 30-100 ng of DNA (Meissner et al. 2008; Boyle et al. 2012; Gu et al. 2010), which maybe more applicable to the quantities of DNA obtained via FACS or LCM. Other studies using genome-wide methylation arrays have corrected for cellular heterogeneity in blood samples using previously determined cell specific DNA methylation marks to deconvolute a mixed cell methylation profile into its component parts relating to the cells of interest (Houseman et al. 2012; Liu et al. 2013), more recently this approach was used to separate neuronal and non-neuronal methylation profiles (Guintivano et al. 2013) from post-mortem brain homogenates.

Another possible limitation relating to post-mortem brain study of DNA methylation, is that factors such as increased post-mortem interval (PMI) and low pH may have an affect on DNA methylation stability. Lowered pH is thought to arise from an increase in tissue lactate as a result of prolonged agonal state and hypoxia (Hynd et al. 2003) and thus depending on the actual cause of death, hypoxia related pH changes in post-mortem tissue may influence measures such as DNA methylation. Though few studies have specifically investigating the effects of pre-, peri-, and post-mortem conditions on DNA methylation status in post-mortem brain tissue, it is believed that DNA methylation status is stable as it involves stable chemical modifications to genomic DNA, which tends to be largely robust to pre-, peri- and post-mortem factors (Ferrer et al. 2008; Pidsley and Mill, 2011). Further still Ernst et al. (2008) found no correlation between DNA methylation levels and pH, while DNA methylation studies in post-mortem brain and artificially induced PMI indicate there to be no effect of PMI on DNA methylation status (Barrachina & Ferrer 2009; Bakulski et al. 2012). However these factors were not accounted for in the studies presented in this thesis due to incomplete data regarding these variables and thus cannot be formally ruled out to have had no effect on the samples tested.

Other unmeasured variables and potential confounders include age, sex, medication, diet and other environmental exposures (toxins, recreational drugs and stress). Whilst age and sex were controlled for in all the analyses, the effects of medication on DNA methylation profiles or gene expression cannot be ruled out along with other unmeasured environmental exposures. For example, though antidepressant treatment was not observed to be associated with SERT methylation levels it was negatively associated with SERT mRNA levels in the small subset of AD patients with depression (Chapter 4).

Better phenotyped samples and clinical data for (both cases and controls) including these unaccounted variables would assist in assessing their effects on DNA methylation and thus identify those variables that need controlling for any confounding effects. Also better characterised post-mortem brains are likely to aid more accurate case-control studies, while more specific histopathological information relating to the actual brain area tested may yield stronger associations with DNA methylation changes based on a semi-quantitative scale of pathology as opposed to absolute disease status. Disease marker associations, such as those presented for Braak staging with DNA methylation in this thesis are limited by the fact that Braak staging only accounts for the neurofibrillary tangle (NFT) pathology, which is secondary to amyloid-β deposition in the brain. Further still both of these neuropathological markers of AD show heterogeneity in the type of structures they form (Cupidi et al. 2010) that adds another layer of complexity. More recent guidelines for the neuropathologic evaluation of AD updating those set out in 1997 (Montine et al. 2012) will no doubt better aid future studies in this direction. Another possible limitation of this thesis worth mentioning is that, while certain brain areas are presumed to be unaffected by AD pathology (i.e. the visual cortex) this is probably an over simplification and though they are likely to be spared of AD pathology such areas may show NFTs and amyloid-β plaques in more advanced disease. Cerebellum, however, is relatively resistant to neuronal damage with little or no tangle formation, and may be a better choice for a brain area that is spared of AD pathology.

Lastly an issue affecting many complex-disease epigenetic studies is the relationship of cause-and-effect. In this thesis DNA methylation was only assessed at the end point of disease in post-mortem tissue after death and while there are observed disease associated differentially methylated signatures these may not necessarily be the cause of disease or predispose an individual to developing AD as these changes in DNA methylation could just as likely be a secondary effect of the neurodegenerative process, or other disease associated factors, such as medication used in treatment that could also lead to altered DNA methylation profiles. Carrying out a parallel assessment of tissues that are not affected by the disease process (such as peripheral lymphocytes or buccal cells) might help distinguish causal from non-causal associations (Petronis 2010).

An ideal study design would be to carry out a prospective mid-to-late life longitudinal study of individuals from a high-risk group or preferably unaffected monozygotic twins (which would control for genetic and early-life environmental influences). Whereby sample collection and recording disease events and other phenotypic changes as well as environmental exposures over a time course would permit detection of the temporal origin and change of disease-

associated DNA methylation variants (Rakyan et al. 2011). Though such a study is not an immediate prospect for AD epigenetic studies, given that the tissue of interest is brain; when and if plausible peripheral tissue epigenetic biomarkers are identified such a study design might be applicable. Furthermore, integration of epigenomic information with genomic, transcriptomic and proteomic data obtained from the same samples will be required to comprehensively assess the role of epigenetic marks such as DNA methylation in AD aetiology.

REFERENCES

Abel, T. & Zukin, R.S. (**2008**) Epigenetic targets of HDAC inhibition in neurodegenerative and psychiatric disorders. *Current Opinion in Pharmacology*, 8(1), 57–64.

Alasaari, J.S., Lagus, M., Ollila, H.M., Toivola, A., Kivimäki, M., Vahtera, J., Kronholm, E., Härmä, M., Puttonen, S. & Paunio, T. (**2012**) Environmental Stress Affects DNA Methylation of a CpG Rich Promoter Region of Serotonin Transporter Gene in a Nurse Cohort. *PLoS ONE*, 7(9), e45813.

Alexopoulos, G.S., Abrams, R.C., Young, R.C. & Shamoian, C.A. (**1988**) Cornell Scale for Depression in Dementia. *Biological Psychiatry*, 23(3), 271–284.

Alonso, A. del C., Zaidi, T., Novak, M., Grundke-Iqbal, I. & Iqbal, K. (**2001**) Hyperphosphorylation induces self-assembly of τ into tangles of paired helical filaments/straight filaments. *Proceedings of the National Academy of Sciences of the United States of America*, 98(12), 6923–6928.

Alzheimer's Association (**2012**) 2012 Alzheimer's disease facts and figures. *Alzheimer's & Dementia*, 8(2), 131–168.

Amir, R.E., Van den Veyver, I.B., Wan, M., Tran, C.Q., Francke, U. & Zoghbi, H.Y. (**1999**) Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nat Genet*, 23(2), 185–188.

Aran, D., Sabato, S. & Hellman, A. (**2013**) DNA methylation of distal regulatory sites characterizes dysregulation of cancer genes. *Genome Biology*, 14(3), R21.

Arango, V., Underwood, M.D., Boldrini, M., Tamir, H., Kassir, S.A., Hsiung, S., Chen, J.J.-X. & Mann, J.J. (**2001**) Serotonin 1A Receptors, Serotonin Transporter Binding and Serotonin Transporter mRNA Expression in the Brainstem of Depressed Suicide Victims. *Neuropsychopharmacology*, 25(6), 892–903.

Assal F, A.M. (**2004**) Association of the serotonin transporter and receptor gene polymorphisms in neuropsychiatric symptoms in alzheimer disease. *Archives of Neurology*, 61(8), 1249–1253.

Athan, L.J. (**2002**) Polymorphisms in the promoter of the human app gene: Functional evaluation and allele frequencies in alzheimer disease. *Archives of Neurology*, 59(11), 1793–1799.

Avramopoulos, D. (**2009**) Genetics of Alzheimer's disease: recent advances. *Genome Medicine*, 1(3), 34.

Avula, R., Rand, A., Black, J.L. & O'Kane, D.J. (**2011**) Simultaneous genotyping of multiple polymorphisms in human serotonin transporter gene and detection of novel allelic variants. *Translational Psychiatry*, 1(8), e32.

Aznar, S. & Knudsen, G.M. (**2011**) Depression and Alzheimer's Disease: Is Stress the Initiating Factor in a Common Neuropathological Cascade? *Journal of Alzheimer's Disease*, 23(2), 177–193.

Babenko, O., Kovalchuk, I. & Metz, G.A. (**2012**) Epigenetic programming of neurodegenerative diseases by an adverse environment. *Brain Research*, 1444(0), 96–111.

Bakulski, K.M., Dolinoy, D.C., Sartor, M.A., Paulson, H.L., Konen, J.R., Lieberman, A.P., Albin, R.L., Hu, H. & Rozek, L.S. (**2012**) Genome-Wide DNA Methylation Differences Between Late-Onset Alzheimer's Disease and Cognitively Normal Controls in Human Frontal Cortex. *Journal of Alzheimer's Disease*, 29(3), 571–588.

Ball, M.P., Li, J.B., Gao, Y., Lee, J.-H., LeProust, E.M., Park, I.-H., Xie, B., Daley, G.Q. & Church, G.M. (**2009**) Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells. *Nat Biotech*, 27(4), 361–368.

Bandeen-Roche, K., Glass, T.A., Bolla, K.I., Todd, Andrew C. & Schwartz, B.S. (**2009**) The Longitudinal Association of Cumulative Lead Dose with Cognitive Function in Community-dwelling Older Adults. *Epidemiology (Cambridge, Mass.)*, 20(6), 831–839.

Barker, D. (**2006**) Commentary: birthweight and coronary heart disease in a historical cohort. *International Journal of Epidemiology*, 35(4), 886–887.

Barker, D., Osmond, C., Kajantie, E. & Eriksson, J.G. (**2009**) Growth and chronic disease: findings in the Helsinki Birth Cohort. *Annals of Human Biology*, 36(5), 445–458.

Barrachina, M & Ferrer, I (**2009**) DNA methylation of Alzheimer disease and tauopathy-related genes in postmortem brain. *J Neuropathol Exp Neurol.*, 68(8), 880–91.

Basha, M. Riyaz, Wei, W., Bakheet, S.A., Benitez, N., Siddiqi, H.K., Ge, Y.-W., Lahiri, D.K. & Zawia, N.H. (**2005**) The Fetal Basis of Amyloidogenesis: Exposure to Lead and Latent Overexpression of Amyloid Precursor Protein and {beta}-Amyloid in the Aging Brain. *J. Neurosci.*, 25(4), 823–829.

Bassett, S.S., Avramopoulos, D. & Fallin, D. (**2002**) Evidence for parent of origin effect in late-onset Alzheimer disease. *American Journal of Medical Genetics*, 114(6), 679–686.

Bassett, S.S., Avramopoulos, D., Perry, R.T., Wiener, H., Watson, B., Go, R.C.P. & Fallin, M.D. (**2006**) Further evidence of a maternal parent-of-origin effect on chromosome 10 in late-onset Alzheimer's disease. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, 141B(5), 537–540.

Bauer, A.P., Leikam, D., Krinner, S., Notka, F., Ludwig, C., Längst, G. & Wagner, R. (**2010**) The impact of intragenic CpG content on gene expression. *Nucleic Acids Research*, 38(12), 3891–3908.

Beach, S.R.H., Brody, G.H., Todorov, A.A., Gunter, T.D. & Philibert, R.A. (**2011**) Methylation at 5HTT Mediates the Impact of Child Sex Abuse on Women's Antisocial Behavior: An Examination of the Iowa Adoptee Sample. *Psychosomatic Medicine*, 73(1), 83–87.

Beach, S.R.H., Brody, G.H., Todorov, A.A., Gunter, T.D. & Philibert, R.A. (**2010**) Methylation at SLC6A4 is linked to family history of child abuse: An examination of the Iowa Adoptee sample. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, 153B(2), 710–713.

Beeri, M.S., Davidson, M., Silverman, J.M., Noy, S., Schmeidler, J. & Goldbourt, U. (**2005**) Relationship between body height and dementia. *The American Journal of Geriatric Psychiatry: Official Journal of the American Association for Geriatric Psychiatry*, 13(2), 116–123.

Bell, R.D., Sagare, A.P., Friedman, A.E., Bedi, G.S., Holtzman, D.M., Deane, R. & Zlokovic, B.V. (**2006**) Transport pathways for clearance of human Alzheimer's amyloid [beta]-peptide and apolipoproteins E and J in the mouse central nervous system. *J Cereb Blood Flow Metab*, 27(5), 909–918.

Benito, E. & Barco, A. (**2010**) CREB's control of intrinsic and synaptic plasticity: implications for CREB-dependent memory models. *Trends in Neurosciences*, 33(5), 230–240.

Berger, S.L. (**2007**) The complex language of chromatin regulation during transcription. *Nature*, 447(7143), 407–412.

Bergsdorf, C., Paliga, K., Kreger, S., Masters, C.L. & Beyreuther, K. (**2000**) Identification of cis-Elements Regulating Exon 15 Splicing of the Amyloid Precursor Protein Pre-mRNA. *Journal of Biological Chemistry*, 275(3), 2046–2056.

Bernstein, B.E., Meissner, A. & Lander, E.S. (**2007**) The Mammalian Epigenome. *Cell*, 128(4), 669–681.

Bertram, L. (**2009**) Alzheimer's disease genetics: Current status and future perspectives. *Alzheimer's and Dementia*, 5(4, Supplement 1), P147.

Bertram, L., McQueen, M.B., Mullin, K., Blacker, D. & Tanzi, R.E. (**2007**) Systematic meta-analyses of Alzheimer disease genetic association studies: the AlzGene database. *Nat Genet*, 39(1), 17–23.

Bettens, K., Sleegers, K. & Van Broeckhoven, C. (**2010**) Current status on Alzheimer disease molecular genetics: from past, to present, to future. *Hum. Mol. Genet.*, ddq142.

Bibikova, M., Barnes, B., Tsan, C., Ho, V., Klotzle, B., Le, J.M., Delano, D., Zhang, Lu, Schroth, G.P., Gunderson, K.L., Fan, J.B. & Shen, R. (**2011**) High density DNA methylation array with single CpG site resolution. *Genomics*, 98(4), 288–295.

Bihaqi, S.W., Huang, H., Wu, J. & Zawia, N.H. (**2011**) Infant Exposure to Lead (Pb) and Epigenetic Modifications in the Aging Primate Brain: Implications for Alzheimer's Disease. *Journal of Alzheimer's Disease*, 27(4), 819–833.

Bird (**2002**) DNA methylation patterns and epigenetic memory. *Genes & Development*, 16(1), 6–21.

Bird, T.D. (**2010**) Alzheimer Disease Overview. URL: http://www.genetests.org.

Bjornsson, H.T., Sigurdsson, M.I., Fallin, M.D., Irizarry, R.A., Aspelund, T., Cui, H., Yu, W., Rongione, M.A., Ekstrom, T.J., Harris, T.B., Launer, L.J., Eiriksdottir, G., Leppert, M.F., Sapienza, C., Gudnason, V. & Feinberg, A.P. (**2008**) Intra-individual Change Over Time in DNA Methylation With Familial Clustering. *JAMA*, 299(24), 2877–2883.

Bolin, C.M., Basha, R., Cox, D., Zawia, N.H., Maloney, B., Lahiri, D.K., & Cardozo-Pelaez, F. (**2006**) Exposure to lead (Pb) and the developmental origin of oxidative DNA damage in the aging brain. *The FASEB Journal*.

Bollati, V., Baccarelli, A., Hou, L., Bonzini, M., Fustinoni, S., Cavallo, D., Byun, H.M., Jiang, J., Marinelli, B., Pesatori, A.C., Bertazzi, P.A. & Yang, A.S. (**2007**) Changes in DNA Methylation Patterns in Subjects Exposed to Low-Dose Benzene. *Cancer Research*, 67(3), 876 –880.

Bollati, V., Schwartz, J., Wright, R., Litonjua, A., Tarantini, L., Suh, H., Sparrow, D., Vokonas, P. & Baccarelli, A. (**2009**) Decline in genomic DNA methylation through aging in a cohort of elderly subjects. *Mechanisms of Ageing and Development*, 130(4), 234–239.

Bollati, V., Galimberti, D., Pergoli, L., Dalla V.E., Barretta, F., Cortini, F., Scarpini, E., Bertazzi, P.A. & Baccarelli, A. (**2011**) DNA methylation in repetitive elements and Alzheimer disease. *Brain, Behavior, and Immunity*, 25(6), 1078–1083.

Booth, M.J., Branco, M.R., Ficz, G., Oxley, D., Krueger, F., Reik, W. & Balasubramanian, S. (**2012**) Quantitative Sequencing of 5-Methylcytosine and 5-Hydroxymethylcytosine at Single-Base Resolution. *Science*.

Borenstein, A.R., Copenhaver, C.I. & Mortimer, J.A. (**2006**) Early-life risk factors for Alzheimer disease. *Alzheimer Disease and Associated Disorders*, 20(1), 63–72.

Borenstein, A.R., Mortimer, J. A., Bowen, J.D., McCormick, W.C., McCurry, S.M., Schellenberg, G.D. & Larson, E.B. (**2001**) Head circumference and incident Alzheimer's disease: Modification by apolipoprotein E. *Neurology*, 57(8), 1453–1460.

Borroni, B., Grassi, M., Agosti, C., Costanzi, C., Archetti, S., Franzoni, S., Caltagirone, C., Di Luca, M., Caimi, L. & Padovani, A. (**2006**) Genetic correlates of behavioral endophenotypes in Alzheimer disease: Role of COMT, 5-HTTLPR and APOE polymorphisms. *Neurobiology of Aging*, 27(11), 1595–1603.

Bouchard, T.J., Lykken, D.T., McGue, M., Segal, N.L. & Tellegen, A. (**1990**) Sources of human psychological differences: the Minnesota Study of Twins Reared Apart. *Science*, 250(4978), 223–228.

Bowen, D.M., Procter, A.W., Mann, D.M.A., Snowden, J.S., Esiri, M.M., Neary, D. & Francis, P. T. (**2008**) Imbalance of a serotonergic system in frontotemporal dementia: implication for pharmacotherapy. *Psychopharmacology*, 196(4), 603–610.

Boyle, P., Clement, K., Gu, H., Smith, Z., Ziller, M., Fostel, J., Holmes, L., Meldrim, J., Kelley, F., Gnirke, A. & Meissner, A. (**2012**) Gel-free multiplexed reduced representation bisulfite sequencing for large-scale DNA methylation profiling. *Genome Biology*, 13(10), R92.

Braak, H. & Braak, E. (**1991**) Neuropathological stageing of Alzheimer-related changes. *Acta Neuropathologica*, 82(4), 239–259.

Bradley, S.L., Dodelzon, K., Sandhu, H.K. & Philibert, R.A. (**2005**) Relationship of serotonin transporter gene polymorphisms and haplotypes to mRNA transcription. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, 136B(1), 58–61.

Branco, M.R., Ficz, G. & Reik, W. (**2012**) Uncovering the role of 5-hydroxymethylcytosine in the epigenome. *Nature Reviews Genetics*, 13(1), 7–13.

Brenet, F., Moh, M., Funk, P., Feierstein, E., Viale, A.J., Socci, N.D. & Scandura, J.M. (**2011**) DNA Methylation of the First Exon Is Tightly Linked to Transcriptional Silencing. *PLoS ONE*, 6(1), e14524.

Brock, B., Basha, R., DiPalma, K., Anderson, A., Harry, G.J., Rice, D.C., Maloney, B., Lahiri, D.K. & Zawia, N.H. (**2008**) Co-localization and Distribution of Cerebral APP and SP1 and its Relationship to Amyloidogenesis. *Journal of Alzheimer's Disease*, 13(1), 71–80.

Brohede, J., Rinde, M., Winblad, Bengt & Graff, C. (**2010**) A DNA Methylation Study of the Amyloid Precursor Protein Gene in Several Brain Regions from Patients with Familial Alzheimer Disease. *Journal of Neurogenetics*, 24(4), 179–181.

Brookmeyer, R., Johnson, E., Ziegler-Graham, K. & Arrighi, H.M. (**2007**) Forecasting the global burden of Alzheimer's disease. *Alzheimer's & Dementia*, 3(3), 186–191.

Brouwers, N., Sleegers, K., Engelborghs, S., Bogaerts, V., Serneels, S., Kamali, K., Corsmit, E., Leenheir, E.D., Martin, J.J., Deyn, P.P.D., Broeckhoven, C.V. & Theuns, J. (**2006**) Genetic risk and transcriptional variability of amyloid precursor protein in Alzheimer's disease. *Brain*, 129(11), 2984–2991.

Brugha, T., Bebbington, P., Tennant, C. & Hurry, J. (**1985**) The List of Threatening Experiences: a subset of 12 life event categories with considerable long-term contextual threat. *Psychological Medicine*, 15(1), 189–194.

Buhot, M.C. (**1997**) Serotonin receptors in cognitive behaviors. *Current Opinion in Neurobiology*, 7(2), 243–254.

Burton, T., Liang, B., Dibrov, A. & Amara, F. (**2002**) Transforming growth factor-β-induced transcription of the Alzheimer β-amyloid precursor protein gene involves interaction between the CTCF-complex and Smads. *Biochemical and Biophysical Research Communications*, 295(3), 713–723.

Cabrejo, L., Guyant-Maréchal, L., Laquerrière, A., Vercelletto, M., Fournière, F.D.L., Thomas-Antérion, C., Verny, C., Letournel, F., Pasquier, F., Vital, A., Checler, F., Frebourg, T., Campion, D. & Hannequin, D. (**2006**) Phenotype associated with APP duplication in five families. *Brain*, 129(11), 2966–2976.

Calin, G.A., Liu, C., Ferracin, M., Hyslop, T., Spizzo, R., Sevignani, C., Fabbri, M., Cimmino, A., Lee, E.J., Wojcik, S.E., Shimizu, M., Tili, E., Rossi, S., Taccioli, C., Pichiorri, F., Liu, X., Zupo, S., Herlea, V., Gramantieri, L., Lanza, G., Alder, H., Rassenti, L., Volinia, S., Schmittgen, T.D., Kipps, T.J., Negrini, M. & Croce, C.M. (**2007**) Ultraconserved Regions Encoding ncRNAs Are Altered in Human Leukemias and Carcinomas. *Cancer Cell*, 12(3), 215–229.

Candore, G., Bulati, M., Caruso, C., Castiglia, L., Colonna-Romano, G., Di Bona, D., Duro, G., Lio, D., Matranga, D., Pellicanò, M., Rizzo, C., Scapagnini, G. & Vasto, S. (**2010**) Inflammation, Cytokines, Immune Response, Apolipoprotein E, Cholesterol, and Oxidative Stress in Alzheimer Disease: Therapeutic Implications. *Rejuvenation Research*, 13(2-3), 301–313.

Casillas, M.A., Jr, Lopatina, N., Andrews, L.G. & Tollefsbol, T.O. (**2003**) Transcriptional control of the DNA methyltransferases is altered in aging and neoplastically-transformed human fibroblasts. *Molecular and Cellular Biochemistry*, 252(1-2), 33–43.

Caspi, A., Sugden, K., Moffitt, T.E., Taylor, A., Craig, I.W., Harrington, H., McClay, J., Mill, Jonathan, Martin, J., Braithwaite, A. & Poulton, R. (**2003**) Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. *Science (New York, N.Y.)*, 301(5631), 386–389.

Cedar, H. & Bergman, Y. (**2009**) Linking DNA methylation and histone modification: patterns and paradigms. *Nat Rev Genet*, 10(5), 295–304.

Champagne, F.A. & Mashoodh, R. (**2009**) Genes in Context: Gene–Environment Interplay and the Origins of Individual Differences in Behavior. *Current Directions in Psychological Science*, 18(3), 127 –131.

Chen, C.P., Eastwood, S.L., Hope, T., McDonald, B., Francis, P.T. & Esiri, M.M. (**2000**) Immunocytochemical study of the dorsal and median raphe nuclei in patients with Alzheimer's disease prospectively assessed for behavioural changes. *Neuropathology and Applied Neurobiology*, 26(4), 347–355.

Chen, C.P., Alder, J.T., Bowen, D.M., Esiri, M.M., McDonald, B., Hope, T., Jobst, K.A. & Francis, P.T. (**1996**) Presynaptic Serotonergic Markers in Community-Acquired Cases of Alzheimer's Disease: Correlations with Depression and Neuroleptic Medication. *Journal of Neurochemistry*, 66(4), 1592–1598.

Chen, H., Dzitoyeva, S. & Manev, H. (**2012**) Effect of aging on 5-hydroxymethylcytosine in the mouse hippocampus. *Restorative Neurology and Neuroscience*, 30(3), 237–245.

Cheon, M.S., Dierssen, M., Kim, S.H. & Lubec, G. (**2008**) Protein expression of BACE1, BACE2 and APP in Down syndrome brains. *Amino Acids*, 35(2), 339–343.

Chodavarapu, R.K., Feng, S., Bernatavichute, Y.V., Chen, P.-Y., Stroud, H., Yu, Y., Hetzel, J.A., Kuo, F., Kim, J., Cokus, S.J., Casero, D., Bernal, M., Huijser, P., Clark, A.T., Krämer, U., Merchant, S.S., Zhang, Xiaoyu, Jacobsen, S.E. & Pellegrini, M. (**2010**) Relationship between nucleosome positioning and DNA methylation. *Nature*, 466(7304), 388–392.

Chouliaras, L., Rutten, B.P.F., Kenis, G., Peerbooms, O., Visser, P.J., Verhey, F., van Os, J., Steinbusch, H.W.M. & van den Hove, D.L.A. (**2010**) Epigenetic regulation in the pathophysiology of Alzheimer's disease. *Progress in Neurobiology*, 90(4), 498–510.

Christensen, B.C., Houseman, E.A., Marsit, C.J., Zheng, S., Wrensch, M.R., Wiemels, J.L., Nelson, H.H., Karagas, M.R., Padbury, J.F., Bueno, R., Sugarbaker, D.J., Yeh, R.-F., Wiencke, J.K. & Kelsey, K.T. (**2009**) Aging and Environmental Exposures Alter Tissue-Specific DNA Methylation Dependent upon CpG Island Context. *PLoS Genet*, 5(8), e1000602.

Clarke, R., Smith, A.D., Jobst, K A, Refsum, H., Sutton, L. & Ueland, P.M. (**1998**) Folate, vitamin B12, and serum total homocysteine levels in confirmed Alzheimer disease. *Archives of Neurology*, 55(11), 1449–1455.

Cohen, M.L., Golde, T.E., Usiak, M.F., Younkin, L.H. & Younkin, S.G. (**1988**) In situ hybridization of nucleus basalis neurons shows increased beta-amyloid mRNA in Alzheimer disease. *Proceedings of the National Academy of Sciences*, 85(4), 1227–1231.

Coolen, M.W., Statham, A.L., Gardiner-Garden, M. & Clark, S.J. (**2007**) Genomic profiling of CpG methylation and allelic specificity using quantitative high-throughput mass spectrometry: critical evaluation and improvements. *Nucleic Acids Research*, 35(18), e119.

Coppedè, F. (**2010**) One-Carbon Metabolism and Alzheimer's Disease: Focus on Epigenetics. , 11(4), 246–260.

Coppus, A., Evenhuis, H., Verberne, G.-J., Visser, F., van Gool, P., Eikelenboom, P. & van Duijin, C. (**2006**) Dementia and mortality in persons with Down's syndrome. *Journal of Intellectual Disability Research: JIDR*, 50(Pt 10), 768–777.

Corder, E., Saunders, A., Strittmatter, W., Schmechel, D.E., Gaskell, P., Small, G., Roses, A.D., Haines, J. & Pericak-Vance, M.A. (**1993**) Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science*, 261(5123), 921–923.

Corder, E.H. & Beaumont, H. (**2007**) Susceptibility groups for Alzheimer's disease (OPTIMA cohort): Integration of gene variants and biochemical factors. *Mechanisms of Ageing and Development*, 128(1), 76–82.

Cupidi, C., Capobianco, R., Goffredo, D., Marcon, G., Ghetti, B., Bugiani, O., Tagliavini, F. & Giaccone, G. (**2010**) Neocortical Variation of Aβ Load in Fully Expressed, Pure Alzheimer's Disease. *Journal of Alzheimer's Disease*, 19(1), 57–68.

Davies, M.N., Volta, M., Pidsley, R., Lunnon, K., Dixit, A., Lovestone, S., Coarfa, C., Harris, R.A., Milosavljevic, A., Troakes, C., Al-Sarraj, S., Dobson, R., Schalkwyk, L.C. & Mill, J. (**2012**) Functional annotation of the human brain methylome identifies tissue-specific epigenetic variation across brain and blood. *Genome Biology*, 13(6), R43.

Davis, C.D. & Uthus, E.O. (**2003**) Dietary Folate and Selenium Affect Dimethylhydrazine-Induced Aberrant Crypt Formation, Global DNA Methylation and One-Carbon Metabolism in Rats. *J. Nutr.*, 133(9), 2907–2914.

DeMattos, R.B., Cirrito, J.R., Parsadanian, M., May, P.C., O'Dell, M.A., Taylor, J.W., Harmony, J.A., Aronow, B.J., Bales, K.R., Paul, S.M. & Holtzman, D.M. (**2004**) ApoE and Clusterin Cooperatively Suppress A[beta] Levels and Deposition: Evidence that ApoE Regulates Extracellular A[beta] Metabolism In Vivo. *Neuron*, 41(2), 193–202.

Van Den Heuvel, C., Thornton, E. & Vink, R. (**2007**) Traumatic brain injury and Alzheimer's disease: a review. *Progress in Brain Research*, 161, 303–316.

Dewji, N.N. & Do, C. (**1996**) Heat shock factor-1 mediates the transcriptional activation of Alzheimer's beta-amyloid precursor protein gene in response to stress. *Brain Research. Molecular Brain Research*, 35(1-2), 325–328.

Doi, A., Park, I.-H., Wen, B., Murakami, P., Aryee, M.J., Irizarry, R., Herb, B., Ladd-Acosta, C., Rho, J., Loewer, S., Miller, J., Schlaeger, T., Daley, G.Q. & Feinberg, A.P. (**2009**) Differential methylation of tissue- and cancer-specific CpG island shores distinguishes human induced pluripotent stem cells, embryonic stem cells and fibroblasts. *Nat Genet*, 41(12), 1350–1353.

Dolinoy, D.C., Weidman, J.R. & Jirtle, R.L. (**2007**) Epigenetic gene regulation: Linking early developmental environment to adult disease. *Reproductive Toxicology*, 23(3), 297–307.

Dover (**2009**) The Barker Hypothesis: How Pediatricans Will Diagnose and Prevent Common Adult-Onset Diseases.

Dunckley, T., Beach, T.G., Ramsey, K.E., Grover, A., Mastroeni, D., Walker, D.G., LaFleur, B.J., Coon, K.D., Brown, K.M., Caselli, R., Kukull, W., Higdon, R., McKeel, D., Morris, J.C., Hulette, C., Schmechel, D., Reiman, E.M., Rogers, J. & Stephan, D.A. (**2006**) Gene expression correlates of neurofibrillary tangles in Alzheimer's disease. *Neurobiology of Aging*, 27(10), 1359–1371.

Durga, J., van Boxtel, M.P.J., Schouten, E.G., Kok, F.J., Jolles, J., Katan, M.B. & Verhoef, P. (**2007**) Effect of 3-year folic acid supplementation on cognitive function in older adults in the FACIT trial: a randomised, double blind, controlled trial. *Lancet*, 369(9557), 208–216.

Dyall, S.C. (**2010**) Amyloid-Beta Peptide, Oxidative Stress and Inflammation in Alzheimer's Disease: Potential Neuroprotective Effects of Omega-3 Polyunsaturated Fatty Acids. *International Journal of Alzheimer's Disease*, 2010, 1–10.

Edwards, C.A. & Ferguson-Smith, A.C. (**2007**) Mechanisms regulating imprinted genes in clusters. *Current Opinion in Cell Biology*, 19(3), 281–289.

Ehrlich, M., Gama-Sosa, M.A., Huang, L.H., Midgett, R.M., Kuo, K.C., McCune, R.A. & Gehrke, C. (**1982**) Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells. *Nucleic Acids Research*, 10(8), 2709–2721.

El-Maarri, O., Walier, M., Behne, F., van Üüm, J., Singer, H., Diaz-Lacava, A., Nüsgen, N., Niemann, B., Watzka, M., Reinsberg, J., van der Ven, H., Wienker, T., Stoffel-Wagner, B., Schwaab, R. & Oldenburg, J. (**2011**) Methylation at Global LINE-1 Repeats in Human Blood Are Affected by Gender but Not by Age or Natural Hormone Cycles. *PLoS ONE*, 6(1), e16252.

Ernst, C., McGowan, P.O., Deleva, V., Meaney, M.J., Szyf, M. & Turecki, G. (**2008**) The effects of pH on DNA methylation state: In vitro and post-mortem brain studies. *Journal of Neuroscience Methods*, 174(1), 123–125.

van Es, M.A. & van den Berg, L.H. (**2009**) Alzheimer's disease beyond APOE. *Nat Genet*, 41(10), 1047–1048.

Espinosa, P.S., Kryscio, R.J., Mendiondo, M.S., Schmitt, F.A., Wekstein, D.R., Markesbery, W.R. & Smith, C.D. (**2006**) Alzheimer's disease and head circumference. *Journal of Alzheimer's Disease*, 9(1), 77–80.

Evans, M.D. & Cooke, M.S. (**2004**) Factors contributing to the outcome of oxidative damage to nucleic acids. *BioEssays*, 26(5), 533–542.

Faghihi, M.A., Modarresi, F., Khalil, A.M., Wood, D.E., Sahagan, B.G., Morgan, T.E., Finch, C.E., St. Laurent III, G., Kenny, P.J. & Wahlestedt, C. (**2008**) Expression of a noncoding RNA is elevated in Alzheimer's disease and drives rapid feed-forward regulation of [beta]-secretase. *Nat Med*, 14(7), 723–730.

Fagiolini, M., Jensen, C.L. & Champagne, F.A. (**2009**) Epigenetic Influences on Brain Development and Plasticity. *Current opinion in neurobiology*, 19(2), 207–212.

Faulkner, G.J., Kimura, Y., Daub, C.O., Wani, S., Plessy, C., Irvine, K.M., Schroder, K., Cloonan, N., Steptoe, A.L., Lassmann, T., Waki, K., Hornig, N., Arakawa, T., Takahashi, H., Kawai, J., Forrest, A.R.R., Suzuki, H., Hayashizaki, Y., Hume, D.A., Orlando, V., Grimmond, S.M. & Carninci, P. (**2009**) The regulated retrotransposon transcriptome of mammalian cells. *Nature Genetics*, 41(5), 563–571.

Feil, R. & Fraga, M.F. (**2012**) Epigenetics and the environment: emerging patterns and implications. *Nat Rev Genet*, 13(2), 97–109.

Ferrer, I., Martinez, A., Boluda, S., Parchi, P. & Barrachina, M. (**2008**) Brain banks: benefits, limitations and cautions concerning the use of post-mortem brain tissue for molecular studies. *Cell and Tissue Banking*, 9(3), 181–194.

Fiala, M. & Veerhuis, R. (**2010**) Biomarkers of inflammation and amyloid-[beta] phagocytosis in patients at risk of Alzheimer disease. *Experimental Gerontology*, 45(1), 57–63.

Ficz, G., Branco, M.R., Seisenberger, S., Krueger, F., Hore, T.A., Marques, C.J., Andrews, S. & Reik, W. (**2011**) Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. *Nature*, 473(7347), 398–402.

Filipowicz, W., Bhattacharyya, S.N. & Sonenberg, N. (**2008**) Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nature Reviews. Genetics*, 9(2), 102–114.

Folstein, M.F., Folstein, S.E. & McHugh, P.R. (**1975**) "Mini-mental state": A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12(3), 189–198.

Forero, D.A., Arboleda, G., Yunis, J.J., Pardo, R. & Arboleda, H. (**2006**) Association study of polymorphisms in LRP1, tau and 5-HTT genes and Alzheimer's disease in a sample of Colombian patients. *Journal of Neural Transmission*, 113(9), 1253–1262.

Forero, D.A., Casadesus, G., Perry, G. & Arboleda, H. (**2006**) Synaptic dysfunction and oxidative stress in Alzheimer's disease: Emerging mechanisms. *Journal of Cellular and Molecular Medicine*, 10(3), 796–805.

Fouse, S.D., Nagarajan, R.P. & Costello, J.F. (**2010**) Genome-scale DNA methylation analysis. *Epigenomics*, 2(1), 105–117.

Fraga, M.F. (**2009**) Genetic and epigenetic regulation of aging. *Current Opinion in Immunology*, 21(4), 446–453.

Fraga, M.F., Ballestar, E., Paz, M.F., Ropero, S., Setien, F., Ballestar, M.L., Heine-Suñer, D., Cigudosa, J.C., Urioste, M., Benitez, J., Boix-Chornet, M., Sanchez-Aguilera, A., Ling, C., Carlsson, E., Poulsen, P., Vaag, A., Stephan, Z., Spector, T.D., Wu, Y.-Z., Plass, C. & Esteller, M. (**2005**) Epigenetic differences arise during the lifetime of monozygotic twins. *Proceedings of the National Academy of Sciences of the United States of America*, 102(30), 10604–10609.

Frodl, T., Reinhold, E., Koutsouleris, N., Donohoe, G., Bondy, B., Reiser, M., Möller, H.-J. & Meisenzahl, E.M. (**2010**) Childhood Stress, Serotonin Transporter Gene and Brain Structures in Major Depression. *Neuropsychopharmacology*, 35(6), 1383–1390.

Frommer, M., McDonald, L.E., Millar, D.S., Collis, C.M., Watt, F., Grigg, G.W., Molloy, P.L. & Paul, C.L. (**1992**) A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proceedings of the National Academy of Sciences of the United States of America*, 89(5), 1827–1831.

Fuks, F., Hurd, P.J., Wolf, D., Nan, Xinsheng, Bird, A.P. & Kouzarides, T. (**2003**) The Methyl-CpG-binding Protein MeCP2 Links DNA Methylation to Histone Methylation. *Journal of Biological Chemistry*, 278(6), 4035–4040.

Fuso, A., Cavallaro, R.A., Zampelli, A., D'Anselmi, F., Piscopo, P., Confaloni, A. & Scarpa, S. (**2007**) γ-Secretase is Differentially Modulated by Alterations of Homocysteine Cycle in Neuroblastoma and Glioblastoma Cells. *Journal of Alzheimer's Disease*, 11(3), 275–290.

Fuso, A., Nicolia, V., Cavallaro, R.A., Ricceri, L., D'Anselmi, F., Coluccia, P., Calamandrei, G. & Scarpa, S. (**2008**) B-vitamin deprivation induces hyperhomocysteinemia and brain S-adenosylhomocysteine, depletes brain S-adenosylmethionine, and enhances PS1 and BACE expression and amyloid-β deposition in mice. *Molecular and Cellular Neuroscience*, 37(4), 731–746.

Fuso, A., Nicolia, V., Pasqualato, A., Fiorenza, M.T., Cavallaro, R.A. & Scarpa, S. (**2011**) Changes in Presenilin 1 gene methylation pattern in diet-induced B vitamin deficiency. *Neurobiology of Aging*, 32(2), 187–199.

Fuso, A. & Scarpa, S. (**2011**) One-carbon metabolism and Alzheimer's disease: is it all a methylation matter? *Neurobiology of Aging*, 32(7), 1192–1195.

Fuso, A., Seminara, L., Cavallaro, R.A., D'Anselmi, F. & Scarpa, S. (**2005**) S-adenosylmethionine/homocysteine cycle alterations modify DNA methylation status with consequent deregulation of PS1 and BACE and beta-amyloid production. *Molecular and Cellular Neuroscience*, 28(1), 195–204.

Garcia-Alloza, M., Gil-Bea, F.J., Diez-Ariza, M., Chen, C.P.L.-H., Francis, P.T., Lasheras, B. & Ramirez, M.J. (**2005**) Cholinergic–serotonergic imbalance contributes to cognitive and behavioral symptoms in Alzheimer's disease. *Neuropsychologia*, 43(3), 442–449.

Gardner, K.L., Hale, M.W., Lightman, S.L., Plotsky, P.M. & Lowry, C.A. (**2009**) Adverse early life experience and social stress during adulthood interact to increase serotonin transporter mRNA expression. *Brain Research*, 1305(0), 47–63.

Gatz, M., Pedersen, N.L., Berg, S., Johansson, B., Johansson, K., Mortimer, James A., Posner, S.F., Viitanen, M., Winblad, Bengt & Ahlbom, A. (**1997**) Heritability for Alzheimer's Disease: The Study of Dementia in Swedish Twins. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 52A(2), M117–M125.

Ge, Y., Ghosh, C., Song, W., Maloney, B. & Lahiri, D.K. (**2004**) Mechanism of promoter activity of the β-amyloid precursor protein gene in different cell lines: identification of a specific 30 bp fragment in the proximal promoter region. *Journal of Neurochemistry*, 90(6), 1432–1444.

Genin, E., Hannequin, D., Wallon, D., Sleegers, K., Hiltunen, M., et al. (**2011**) APOE and Alzheimer Disease: A Major Gene with Semi-Dominant Inheritance. *Molecular psychiatry*, 16(9), 903–907.

Giannakopoulos, P., Herrmann, F.R., Bussière, T., Bouras, C., Kövari, E., Perl, D.P., Morrison, J.H., Gold, G. & Hof, P.R. (**2003**) Tangle and neuron numbers, but not amyloid load, predict cognitive status in Alzheimer's disease. *Neurology*, 60(9), 1495–1500.

Giardine, B., Riemer, C., Hardison, R.C., Burhans, R., Elnitski, L., Shah, P., Zhang, Yi, Blankenberg, D., Albert, I., Taylor, J., Miller, W., Kent, W.J. & Nekrutenko, A. (**2005**) Galaxy: A platform for interactive large-scale genome analysis. *Genome Research*, 15(10), 1451–1455.

Gibbs, J.R., van der Brug, M.P., Hernandez, D.G., Traynor, B.J., Nalls, M.A., Lai, S.-L., Arepalli, S., Dillman, A., Rafferty, I.P., Troncoso, J., Johnson, R., Zielke, H.R., Ferrucci, L., Longo, D.L., Cookson, M.R. & Singleton, A.B. (**2010**) Abundant Quantitative Trait Loci Exist for DNA Methylation and Gene Expression in Human Brain. *PLoS Genet*, 6(5), e1000952.

Gillespie, N.A., Whitfield, J.B., Williams, Ben, Heath, A.C. & Martin, N.G. (**2005**) The relationship between stressful life events, the serotonin transporter (5-HTTLPR) genotype and major depression. *Psychological Medicine*, 35(1), 101–111.

Gilley, D.W., Wilson, R.S., Bienias, J.L., Bennett, D.A. & Evans, D.A. (**2004**) Predictors of Depressive Symptoms in Persons With Alzheimer's Disease. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 59(2), P75–P83.

Globisch, D., Münzel, M., Müller, M., Michalakis, S., Wagner, M., Koch, S., Brückl, T., Biel, M. & Carell, T. (**2010**) Tissue Distribution of 5-Hydroxymethylcytosine and Search for Active Demethylation Intermediates. *PLoS ONE*, 5(12), e15367.

Gluckman, P.D., Hanson, M.A. & Beedle, A.S. (**2007**) Early life events and their consequences for later disease: A life history and evolutionary perspective. *American Journal of Human Biology*, 19(1), 1–19.

Goate, A., Chartier-Harlin, M.-C., Mullan, M., Brown, J., Crawford, F., Fidani, L., Giuffra, L., Haynes, A., Irving, N., James, L., Mant, R., Newton, P., Rooke, K., Roques, P., Talbot, C., Pericak-Vance, M., Roses, A., Williamson, R., Rossor, M., Owen, M. & Hardy, J. (**1991**) Segregation of a missense mutation in the amyloid precursor protein gene with familial Alzheimer's disease. *Nature*, 349(6311), 704–706.

Goedert, M. & Jakes, R. (**2005**) Mutations causing neurodegenerative tauopathies. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, 1739(2-3), 240–250.

Golde, T.E., Estus, S., Usiak, M., Younkin, L.H. & Younkin, S.G. (**1990**) Expression of β amyloid protein precursor mRNAs: Recognition of a novel alternatively spliced form and quantitation in alzheimer's disease using PCR. *Neuron*, 4(2), 253–267.

Graeber, M.B., Kösel, S., Egensperger, R., Banati, R.B., Müller, U., Bise, K., Hoff, P., Möller, H.J., Fujisawa, K. & Mehraein, P. (**1997**) Rediscovery of the case described by Alois Alzheimer in 1911: historical, histological and molecular genetic analysis. *Neurogenetics*, 1(1), 73–80.

Groom, A., Elliott, H.R., Embleton, N.D. & Relton, C.L. (**2011**) Epigenetics and child health: basic principles. *Archives of Disease in Childhood*, 96(9), 863–869.

Grünblatt, E., Zehetmayer, S., Bartl, J., Löffler, C., Wichart, I., Rainer, M.K., Jungwirth, S., Bauer, P., Danielczyk, W., Tragl, K.-H., Riederer, P. & Fischer, P. (**2009**) Genetic risk factors and markers for Alzheimer's disease and/or depression in the VITA study. *Journal of Psychiatric Research*, 43(3), 298–308.

Grupe, A., Abraham, R., Li, Y., Rowland, C., Hollingworth, P., Morgan, A., Jehu, L., Segurado, R., Stone, D., Schadt, E., Karnoub, M., Nowotny, P., Tacey, K., Catanese, J., Sninsky, J., Brayne, C., Rubinsztein, D., Gill, M., Lawlor, B., Lovestone, S., Holmans, P., O'Donovan, M., Morris, J.C., Thal, L., Goate, A., Owen, M.J. & Williams, J. (**2007**) Evidence for novel susceptibility genes for late-onset Alzheimer's disease from a genome-wide association study of putative functional variants. *Human Molecular Genetics*, 16(8), 865–873.

Gu, H., Bock, C., Mikkelsen, T.S., Jäger, N., Smith, Z.D., Tomazou, E., Gnirke, A., Lander, E.S. & Meissner, A. (**2010**) Genome-scale DNA methylation mapping of clinical samples at single-nucleotide resolution. *Nature Methods*, 7(2), 133–136.

Gu, X., Sun, J., Li, S., Wu, Xiangmei & Li, Liang (**2013**) Oxidative stress induces DNA demethylation and histone acetylation in SH-SY5Y cells: potential epigenetic mechanisms in gene transcription in Aβ production. *Neurobiology of Aging*, 34(4), 1069–1079.

Guerreiro, R., Wojtas, A., Bras, J., Carrasquillo, M., Rogaeva, E., Majounie, E., Cruchaga, C., Sassi, C., Kauwe, J.S.K., Younkin, S., Hazrati, L., Collinge, J., Pocock, J., Lashley, T., Williams, J., Lambert, J.-C., Amouyel, P., Goate, A., Rademakers, R., Morgan, K., Powell, J., St. George-Hyslop, P., Singleton, A. & Hardy, J. (**2013**) TREM2 Variants in Alzheimer's Disease. *New England Journal of Medicine*, 368(2), 117–127.

Guintivano, J., Aryee, M.J. & Kaminsky, Z.A. (**2013**) A cell epigenotype specific model for the correction of brain cellular heterogeneity bias and its application to age, brain region and major depression. *Epigenetics*, 8(3), 290–302.

Guo, J.U., Su, Y., Zhong, C., Ming, G. & Song, H. (**2011**) Hydroxylation of 5-Methylcytosine by TET1 Promotes Active DNA Demethylation in the Adult Brain. *Cell*, 145(3), 423–434.

Guttula, S.V, Allam, A. & Gumpeny, R.S. (**2012**) Analyzing Microarray Data of Alzheimer's Using Cluster Analysis to Identify the Biomarker Genes. *International Journal of Alzheimer's Disease*, 2012, 1–5.

Guyant-Marechal, L., Rovelet-Lecrux, A., Goumidi, L., Cousin, E., Hannequin, D., Raux, G., Penet, C., Ricard, S., Mace, S., Amouyel, P., Deleuze, J.-F., Frebourg, T., Brice, A., Lambert, J.-C. & Campion, D. (**2007**) Variations in the APP gene promoter region and risk of Alzheimer disease. *Neurology*, 68(9), 684–687.

Haan, M.N., Miller, J.W., Aiello, A.E., Whitmer, R.A., Jagust, W.J., Mungas, D.M., Allen, L.H. & Green, R. (**2007**) Homocysteine, B vitamins, and the incidence of dementia and cognitive impairment: results from the Sacramento Area Latino Study on Aging. *Am J Clin Nutr*, 85(2), 511–517.

Hagen, G., Dennig, J., Preiß, A., Beato, M. & Suske, G. (**1995**) Functional Analyses of the Transcription Factor Sp4 Reveal Properties Distinct from Sp1 and Sp3. *Journal of Biological Chemistry*, 270(42), 24989–24994.

Hakansson, K., Rovio, S., Helkala, E.-L., Vilska, A.-R., Winblad, B., Soininen, H., Nissinen, A., Mohammed, A.H. & Kivipelto, M. (**2009**) Association between mid-life marital status and cognitive function in later life: population based cohort study. *BMJ : British Medical Journal*, 339.

Halliwell, B. (**2006**) Oxidative stress and neurodegeneration: where are we now? *Journal of Neurochemistry*, 97(6), 1634–1658.

Hampel, H., Bürger, K., Teipel, S.J., Bokde, A.L.W., Zetterberg, H. & Blennow, K. (**2008**) Core candidate neurochemical and imaging biomarkers of Alzheimer's disease. *Alzheimer's & Dementia*, 4(1), 38–48.

Hardy, J. (**2006**) A Hundred Years of Alzheimer's Disease Research. *Neuron*, 52(1), 3–13.

Hardy, J. & Selkoe, D.J. (**2002**) The Amyloid Hypothesis of Alzheimer's Disease: Progress and Problems on the Road to Therapeutics. *Science*, 297(5580), 353–356.

Harold, D., Abraham, R., Hollingworth, P., Sims, R., Gerrish, A., et al. (**2009**) Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. *Nat Genet*, 41(10), 1088–1093.

Hartmann, T., Bergsdorf, C., Sandbrink, R., Tienari, P.J., Multhaup, G., Ida, N., Bieger, S., Dyrks, T., Weidemann, A., Masters, Colin L. & Beyreuther, K. (**1996**) Alzheimer's Disease βA4 Protein Release and Amyloid Precursor Protein Sorting Are Regulated by Alternative Splicing. *Journal of Biological Chemistry*, 271(22), 13208–13214.

Hayden, K.M., Norton, M.C., Darcey, D., Ostbye, T., Zandi, P.P., Breitner, J.C.S. & Welsh-Bohmer, K.A. (**2010**) Occupational exposure to pesticides increases the risk of incident AD. *Neurology*, 74(19), 1524–1530.

He, Y.-F., Li, B.-Z., Li, Z., Liu, P., Wang, Y., Tang, Q., Ding, J., Jia, Y., Chen, Z., Li, Lin, Sun, Yan, Li, Xiuxue, Dai, Q., Song, C.-X., Zhang, K., He, C. & Xu, G.-L. (**2011**) Tet-Mediated Formation of 5-Carboxylcytosine and Its Excision by TDG in Mammalian DNA. *Science*, 333(6047), 1303–1307.

Heils, A., Teufel, A., Petri, S., Stöber, G., Riederer, P., Bengel, D., & Lesch, K.P. (**1996**) Allelic Variation of Human Serotonin Transporter Gene Expression. *Journal of Neurochemistry*, 66(6), 2621–2624.

Heinz, A., Jones, D.W., Mazzanti, C., Goldman, D., Ragan, P., Hommer, D., Linnoila, M. & Weinberger, D.R. (**2000**) A relationship between serotonin transporter genotype and in vivo protein expression and alcohol neurotoxicity. *Biological Psychiatry*, 47(7), 643–649.

Henderson, I.R., Chan, S.R., Cao, X., Johnson, L. & Jacobsen, S.E. (**2010**) Accurate sodium bisulfite sequencing in plants. *Epigenetics : official journal of the DNA Methylation Society*, 5(1), 47–49.

Heyn, H., Li, N., Ferreira, H.J., Moran, S., Pisano, D.G., Gomez, A., Diez, J., Sanchez-Mut, J.V., Setien, F., Carmona, F.J., Puca, A.A., Sayols, S., Pujana, M.A., Serra-Musach, J., Iglesias-Platas, I., Formiga, F., Fernandez, A.F., Fraga, M.F., Heath, S.C., Valencia, A., Gut, I.G., Wang, Jun & Esteller, M. (**2012**) Distinct DNA methylomes of newborns and centenarians. *Proceedings of the National Academy of Sciences*, 109(26), 10522–10527.

Higgins, G.A., Lewis, D.A., Bahmanyar, S., Goldgaber, D., Gajdusek, D.C., Young, W.G., Morrison, J.H. & Wilson, M.C. (**1988**) Differential regulation of amyloid-beta-protein mRNA expression within hippocampal neuronal subpopulations in Alzheimer disease. *Proceedings of the National Academy of Sciences*, 85(4), 1297–1301.

Hollants, S., Redeker, E.J.W. & Matthijs, G. (**2012**) Microfluidic Amplification as a Tool for Massive Parallel Sequencing of Familial Hypercholesterolemia Genes. *Clinical Chemistry*.

Hollingworth, P., Harold, D., Sims, R., Gerrish, A., Lambert, J.-C., et al. (**2011**) Common variants at ABCA7, MS4A6A/MS4A4E, EPHA1, CD33 and CD2AP are associated with Alzheimer's disease. *Nature Genetics*, 43(5), 429–435.

Hooli, B.V., Mohapatra, G., Mattheisen, M., Parrado, A.R., Roehr, J.T., Shen, Y., Gusella, J.F., Moir, R., Saunders, A.J., Lange, C., Tanzi, R.E. & Bertram, L. (**2012**) Role of common and rare APP DNA sequence variants in Alzheimer disease. *Neurology*, 78(16), 1250–1257.

Houseman, E.A, Accomando, W.P., Koestler, D.C., Christensen, B.C., Marsit, C.J., Nelson, H.H., Wiencke, J.K. & Kelsey, K.T. (**2012**) DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics*, 13(1), 86.

Hu, M., Retz, W., Baader, M., Pesold, B., Adler, G., Henn, F.A., Rösler, M. & Thome, J. (**2000**) Promoter polymorphism of the 5-HT transporter and Alzheimer's disease. *Neuroscience Letters*, 294(1), 63–65.

Huang, Y., Pastor, W.A., Shen, Y., Tahiliani, M., Liu, D.R. & Rao, A. (**2010**) The Behaviour of 5-Hydroxymethylcytosine in Bisulfite Sequencing. *PLoS ONE*, 5(1), e8888.

Huezo-Diaz, P., Uher, R., Smith, R., Rietschel, M., Henigsberg, N., Marušič, A., Mors, O., Maier, W., Hauser, J., Souery, D., Placentino, A., Zobel, A., Larsen, E.R., Czerski, P.M., Gupta, B., Hoda, F., Perroud, N., Farmer, A., Craig, I., Aitchison, K.J. & McGuffin, P. (**2009**) Moderation of antidepressant response by the serotonin transporter gene. *The British Journal of Psychiatry*, 195(1), 30–38.

Hynd, M.R., Lewohl, J.M., Scott, H.L. & Dodd, P.R. (**2003**) Biochemical and molecular studies using human autopsy brain tissue. *Journal of Neurochemistry*, 85(3), 543–562.

Illingworth, R.S. & Bird, A.P. (**2009**) CpG islands - "A rough guide." *FEBS Letters*, 583(11), 1713–1720.

Illingworth, R.S., Gruenewald-Schneider, U., Webb, S., Kerr, A.R.W., James, K.D., Turner, D.J., Smith, C., Harrison, D.J., Andrews, R. & Bird, A.P. (**2010**) Orphan CpG Islands Identify Numerous Conserved Promoters in the Mammalian Genome. *PLoS Genet*, 6(9), e1001134.

Irizarry, R.A., Ladd-Acosta, C., Wen, B., Wu, Z., Montano, C., Onyango, P., Cui, H., Gabo, K., Rongione, M., Webster, M., Ji, H., Potash, J.B., Sabunciyan, S. & Feinberg, A.P. (**2009**) The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. *Nat Genet*, 41(2), 178–186.

Ito, S., D'Alessio, A.C., Taranova, O.V., Hong, K., Sowers, L.C. & Zhang, Yi (**2010**) Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature*, advance online publication.

Jenuwein, T. & Allis, C.D. (**2001**) Translating the histone code. *Science (New York, N.Y.)*, 293(5532), 1074–1080.

Jin, S.-G., Kadam, S. & Pfeifer, G.P. (**2010**) Examination of the specificity of DNA methylation profiling techniques towards 5-methylcytosine and 5-hydroxymethylcytosine. *Nucl. Acids Res.*, gkq223.

Jin, S.-G., Wu, Xiwei, Li, A.X. & Pfeifer, G.P. (**2011**) Genomic Mapping of 5-Hydroxymethylcytosine in the Human Brain. *Nucleic Acids Research*, 39(12), 5015–5024.

Johnstone, S.E. & Baylin, S.B. (**2010**) Stress and the epigenetic landscape: a link to the pathobiology of human diseases? *Nat Rev Genet*, 11(11), 806–812.

Jones et al. (**2010**) Genetic evidence implicates the immune system and cholesterol metabolism in the aetiology of Alzheimer's disease.

Jones, P.A. & Baylin, S.B. (**2007**) The epigenomics of cancer. *Cell*, 128(4), 683–692.

Jones, P.A. & Takai, D. (**2001**) The Role of DNA Methylation in Mammalian Epigenetics. *Science*, 293(5532), 1068–1070.

Kadonaga, J.T., Carner, K.R., Masiarz, F.R. & Tjian, R. (**1987**) Isolation of cDNA encoding transcription factor Sp1 and functional analysis of the DNA binding domain. *Cell*, 51(6), 1079–1090.

Kageyama, M., Hiraoka, M. & Kagawa, Y. (**2008**) Relationship between genetic polymorphism, serum folate and homocysteine in Alzheimer's disease. *Asia-Pacific Journal of Public Health / Asia-Pacific Academic Consortium for Public Health*, 20 Suppl, 111–117.

Kálmán, J., Kitajka, K., Pákáski, M., Zvara, A., Juhász, A., Vincze, G., Janka, Z. & Puskás, L.G. (**2005**) Gene expression profile analysis of lymphocytes from Alzheimer's patients. *Psychiatric Genetics*, 15(1), 1–6.

Karimi, M., Johansson, S., Stach, D., Corcoran, M., Grandér, D., Schalling, M., Bakalkin, G., Lyko, F., Larsson, C. & Ekström, T.J. (**2006**) LUMA (LUminometric Methylation Assay)--A high throughput method to the analysis of genomic DNA methylation. *Experimental Cell Research*, 312(11), 1989–1995.

Kasuga, K., Shimohata, T., Nishimura, A., Shiga, A., Mizuguchi, T., Tokunaga, J., Ohno, T., Miyashita, A., Kuwano, R., Matsumoto, N., Onodera, O., Nishizawa, M. & Ikeuchi, T. (**2009**) Identification of independent APP locus duplication in Japanese patients with early-onset Alzheimer disease. *Journal of Neurology, Neurosurgery & Psychiatry*, 80(9), 1050 –1052.

Kazazian, H.H. (**2004**) Mobile Elements: Drivers of Genome Evolution. *Science*, 303(5664), 1626 –1632.

Kennedy, J.L., Farrer, L.A., Andreasen, N.C., Mayeux, R. & St George-Hyslop, P. (**2003**) The Genetics of Adult-Onset Neuropsychiatric Disease: Complexities and Conundra? *Science*, 302(5646), 822–826.

Kepe, V., Barrio, J.R., Huang, S.-C., Ercoli, L., Siddarth, P., Shoghi-Jadid, K., Cole, G.M., Satyamurthy, N., Cummings, J.L., Small, G.W. & Phelps, M.E. (**2006**) Serotonin 1A receptors in the living brain of Alzheimer's disease patients. *Proceedings of the National Academy of Sciences of the United States of America*, 103(3), 702–707.

Khare, T., Pai, S., Koncevicius, K., Pal, M., Kriukiene, E., Liutkeviciute, Z., Irimia, M., Jia, P., Ptak, C., Xia, M., Tice, R., Tochigi, M., Moréra, S., Nazarians, A., Belsham, D., Wong, A.H.C., Blencowe, B.J., Wang, S.C., Kapranov, P., Kustra, R., Labrie, V., Klimasauskas, S. & Petronis, A. (**2012**) 5-hmC in the brain is abundant in synaptic genes and shows differences at the exon-intron boundary. *Nature Structural & Molecular Biology*, 19(10), 1037–1043.

Kim, J., Kim, J.-Y.,& Issa, J.-P.J. (**2009**) Aging and DNA Methylation. *Current Chemical Biology*, 3(1), 1–9.

Kinnally, E.L., Capitanio, J.P., Leibel, R., Deng, L., LeDuc, C., Haghighi, F. & Mann, J.J. (**2010**) Epigenetic regulation of serotonin transporter expression and behavior in infant rhesus macaques. *Genes, Brain and Behavior*, 9(6), 575–582.

Kinnally, E.L., Feinberg, C., Kim, D., Ferguson, K., Leibel, R., Coplan, J.D. & Mann, J.J. (**2011**) DNA methylation as a risk factor in the effects of early life stress. *Brain, Behavior, and Immunity*, 25(8), 1548–1553.

Kinney, S.M., Chin, H.G., Vaisvila, R., Bitinaite, J., Zheng, Y., Estève, P.-O., Feng, S., Stroud, H., Jacobsen, S.E. & Pradhan, S. (**2011**) Tissue-specific Distribution and Dynamic Changes of 5-Hydroxymethylcytosine in Mammalian Genomes. *Journal of Biological Chemistry*, 286(28), 24685 –24693.

Komori, H.K., LaMere, S.A., Torkamani, A., Hart, G.T., Kotsopoulos, S., Warner, J., Samuels, M.L., Olson, J., Head, S.R., Ordoukhanian, P., Lee, P.L., Link, D.R. & Salomon, D.R. (**2011**) Application of microdroplet PCR for large-scale targeted bisulfite sequencing. *Genome Research*.

König, G, Mönning, U., Czech, C., Prior, R., Banati, R, Schreiter-Gasser, U., Bauer, J., Masters, C L & Beyreuther, K (**1992**) Identification and differential expression of a novel alternative splice isoform of the beta A4 amyloid precursor protein (APP) mRNA in leukocytes and brain microglial cells. *The Journal of Biological Chemistry*, 267(15), 10804–10809.

Koo, E.H., Sisodia, S.S., Cork, L.C., Unterbeck, A., Bayney, R.M. & Price, D.L. (**1990**) Differential expression of amyloid precursor protein mRNAs in cases of Alzheimer's disease and in aged nonhuman primates. *Neuron*, 4(1), 97–104.

Kraus, T.F.J., Globisch, D., Wagner, M., Eigenbrod, S., Widmann, D., Münzel, M., Müller, M., Pfaffeneder, T., Hackner, B., Feiden, W., Schüller, U., Carell, T. & Kretzschmar, H.A. (**2012**) Low values of 5-hydroxymethylcytosine (5hmC), the "sixth base," are associated with anaplasia in human brain tumors. *International Journal of Cancer*, n/a–n/a.

Kriaucionis, S. & Heintz, N. (**2009**) The Nuclear DNA Base 5-Hydroxymethylcytosine Is Present in Purkinje Neurons and the Brain. *Science*, 324(5929), 929–930.

Kronenberg, G., Colla, M. & Endres, M. (**2009**) Folic Acid, Neurodegenerative and Neuropsychiatric Disease. *Current Molecular Medicine*, 9(3), 315–323.

Krueger, F. & Andrews, S.R. (**2011**) Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics*, 27(11), 1571–1572.

Krueger, F., Kreck, B., Franke, A. & Andrews, S.R. (**2012**) DNA methylome analysis using short bisulfite sequencing data. *Nature Methods*, 9(2), 145–151.

Kruman, I.I., Kumaravel, T.S., Lohani, A., Pedersen, W.A., Cutler, R.G., Kruman, Y., Haughey, N., Lee, J., Evans, M. & Mattson, M.P. (**2002**) Folic Acid Deficiency and Homocysteine Impair DNA Repair in Hippocampal Neurons and Sensitize Them to Amyloid Toxicity in Experimental Models of Alzheimer's Disease. *J. Neurosci.*, 22(5), 1752–1762.

Kunugi, H., Ueki, Akira, Otsuka, M., Isse, K., Hirasawa, H., Kato, N., Nabika, T., Kobayashi, S. & Nanko, S. (**2000**) Alzheimer's disease and 5-HTTLPR polymorphism of the serotonin transporter gene: No evidence for an association. *American Journal of Medical Genetics*, 96(3), 307–309.

Ladd-Acosta, C., Pevsner, J., Sabunciyan, S., Yolken, R.H., Webster, M.J., Dinkins, T., Callinan, P.A., Fan, J.-B., Potash, J.B. & Feinberg, A.P. (**2007**) DNA Methylation Signatures within the Human Brain. *The American Journal of Human Genetics*, 81(6), 1304–1315.

LaFerla, F.M. (**2002**) Calcium dyshomeostasis and intracellular signalling in alzheimer's disease. *Nature Reviews Neuroscience*, 3(11), 862–872.

Lahiri, D.K., Nall, C. & Ge, Y.W. (**1999**) Promoter activity of the beta-amyloid precursor protein gene is negatively modulated by an upstream regulatory element. *Brain Research. Molecular Brain Research*, 71(1), 32–41.

Lahiri, D.K. (**2004**) Functional Characterization of Amyloid β Precursor Protein Regulatory Elements: Rationale for the Identification of Genetic Polymorphism. *Annals of the New York Academy of Sciences*, 1030(1), 282–288.

Lahiri, D.K. & Ge, Y.-W. (**2004**) Role of the APP Promoter in Alzheimer's Disease: Cell Type-Specific Expression of the β-Amyloid Precursor Protein. *Annals of the New York Academy of Sciences*, 1030(1), 310–316.

Lahiri, D.K., Ge, Y.-W. & Maloney, B. (**2005**) Characterization of the APP proximal promoter and 5'-untranslated regions: identification of cell-type specific domains and implications in APP gene expression and Alzheimer's disease. *The FASEB Journal*.

Lahiri, D.K. & Maloney, B. (**2010**) The "LEARn" (Latent Early-life Associated Regulation) model integrates environmental risk factors and the developmental basis of Alzheimer's disease, and proposes remedial steps. *Experimental Gerontology*, 45(4), 291–296.

Lahiri, D.K. & Robakis, N.K. (**1991**) The promoter activity of the gene encoding Alzheimer β-amyloid precursor protein (APP) is regulated by two blocks of upstream sequences. *Molecular Brain Research*, 9(3), 253–257.

Lai, M.K., Tsang, S.W., Alder, J.T., Keene, J., Hope, T., Esiri, M.M., Francis, P.T. & Chen, C.P. (**2005**) Loss of serotonin 5-HT2A receptors in the postmortem temporal cortex correlates with rate of cognitive decline in Alzheimer's disease. *Psychopharmacology*, 179(3), 673–677.

Lai, M.K.P., Tsang, S.W., Esiri, M.M., Francis, P.T., Wong, P.T.-H. & Chen, C.P. (**2011**) Differential involvement of hippocampal serotonin1A receptors and re-uptake sites in non-cognitive behaviors of Alzheimer's disease. *Psychopharmacology*, 213(2-3), 431–439.

Lai, M.K.P., Tsang, S.W.Y., Francis, P.T., Keene, J., Hope, T., Esiri, M.M., Spence, I. & Chen, C.P (**2002**) Postmortem serotoninergic correlates of cognitive decline in Alzheimer's disease. *Neuroreport*, 13(9), 1175–1178.

Lambert, J.C., Grenier-Boley, B., Chouraki, V., Heath, S., Zelenika, D., Fievet, N., Hannequin, D., Pasquier, F., Hanon, O., Brice, A., Epelbaum, J., Berr, C., Dartigues, J.-F., Tzourio, C., Campion, D., Lathrop, M. & Amouyel, P. (**2010**) Implication of the Immune System in Alzheimer's Disease: Evidence from Genome-Wide Pathway Analysis. *Journal of Alzheimer's Disease*, [Epub ahead of print].

Lambert, J.C., Heath, S., Even, G., Campion, D., Sleegers, K., Hiltunen, M., Combarros, O., Zelenika, D., Bullido, M.J., Tavernier, B., Letenneur, L., Bettens, K., Berr, C., Pasquier, F., Fievet, N., Barberger-Gateau, P., Engelborghs, S., De Deyn, P., Mateo, I., Franck, A., Helisalmi, S., Porcellini, E., Hanon, O., de Pancorbo, M.M., Lendon, C., Dufouil, C., Jaillard, C., Leveillard, T., Alvarez, V., Bosco, P., Mancuso, Michelangelo, Panza, F., Nacmias, B., Bossu, P., Piccardi, P., Annoni, G., Seripa, D., Galimberti, Daniela, Hannequin, Didier, Licastro, F., Soininen, H., Ritchie, K., Blanche, H., Dartigues, J.-F., Tzourio, C., Gut, I., Van Broeckhoven, C., Alperovitch, A., Lathrop, M. & Amouyel, P. (**2009**) Genome-wide association study identifies variants at CLU and CR1 associated with Alzheimer's disease. *Nat Genet*, 41(10), 1094–1099.

Lande-Diner, L., Zhang, J., Ben-Porath, I., Amariglio, N., Keshet, I., Hecht, M., Azuara, V., Fisher, A.G., Rechavi, G. & Cedar, H. (**2007**) Role of DNA Methylation in Stable Gene Repression. *Journal of Biological Chemistry*, 282(16), 12194–12200.

Lasky-Su, J.A., Faraone, S.V., Glatt, S.J. & Tsuang, M.T. (**2005**) Meta-analysis of the association between two polymorphisms in the serotonin transporter gene and affective disorders. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, 133B(1), 110–115.

Laurent, L., Wong, E., Li, G., Huynh, T., Tsirigos, A., Ong, C.T., Low, H.M., Kin Sung, K.W., Rigoutsos, I., Loring, J. & Wei, C.-L. (**2010**) Dynamic changes in the human methylome during differentiation. *Genome Research*, 20(3), 320–331.

Le, T., Kim, K.-P., Fan, G. & Faull, K.F. (**2011**) A sensitive mass spectrometry method for simultaneous quantification of DNA methylation and hydroxymethylation levels in biological samples. *Analytical Biochemistry*, 412(2), 203–209.

Ledoux, S., Nalbantoglu, J. & Cashman, N.R. (**1994**) Amyloid precursor protein gene expression in neural cell lines: influence of DNA cytosine methylation. *Molecular Brain Research*, 24(1–4), 140–144.

Lee, K.S., Cheong, H.-K., Eom, J.-S., Jung, H.S., Oh, B.H. & Hong, C.H. (**2009**) Cognitive decline is associated with nutritional risk in subjects with small head circumference (HC). *Archives of Gerontology and Geriatrics*, In Press, Corrected Proof.

Lesch, K.P., Bengel, D., Heils, A., Sabol, S.Z., Greenberg, B.D., Petri, S, Benjamin, J., Müller, C.R., Hamer, D.H. & Murphy, D.L. (**1996**) Association of anxiety-related traits with a polymorphism in the serotonin transporter gene regulatory region. *Science (New York, N.Y.)*, 274(5292), 1527–1531.

Levenson, J.M., Roth, T.L., Lubin, F.D., Miller, C.A., Huang, I.-C., Desai, P., Malone, L.M. & Sweatt, J.D. (**2006**) Evidence That DNA (Cytosine-5) Methyltransferase Regulates Synaptic Plasticity in the Hippocampus. *Journal of Biological Chemistry*, 281(23), 15763–15773.

Levy-Lahad, E., Wasco, W., Poorkaj, P., Romano, D., Oshima, J., Pettingell, W., Yu, C., Jondro, P., Schmidt, S., Wang, K. et al. (**1995**) Candidate gene for the chromosome 1 familial Alzheimer's disease locus. *Science*, 269(5226), 973–977.

Lewis, D.A., Higgins, G.A., Young, W.G., Goldgaber, D., Gajdusek, D.C., Wilson, M.C. & Morrison, J.H. (**1988**) Distribution of precursor amyloid-beta-protein messenger RNA in human cerebral cortex: relationship to neurofibrillary tangles and neuritic plaques. *Proceedings of the National Academy of Sciences*, 85(5), 1691–1695.

Li, K., Singh, A., Crooks, D.R., Dai, X., Cong, Z., Pan, L., Ha, D. & Rouault, T.A. (**2010**) Expression of Human Frataxin Is Regulated by Transcription Factors SRF and TFAP2. *PLoS ONE*, 5(8), e12286.

Li, L.-C. & Dahiya, R. (**2002**) MethPrimer: designing primers for methylation PCRs. *Bioinformatics (Oxford, England)*, 18(11), 1427–1431.

Li, T., Holmes, C, Sham, P.C., Vallada, H., Birkett, J., Kirov, G., Lesch, K P, Powell, J, Lovestone, S & Collier, D. (**1997**) Allelic functional variation of serotonin transporter expression is a susceptibility factor for late onset Alzheimer's disease. *Neuroreport*, 8(3), 683–686.

Li, Y.-Y., Chen, T., Wan, Y. & Xu, S. (**2012**) Lead exposure in pheochromocytoma cells induces persistent changes in amyloid precursor protein gene methylation patterns. *Environmental Toxicology*, 27(8), 495–502.

Liang, W.S., Dunckley, T., Beach, T.G., Grover, A., Mastroeni, D., Ramsey, K., Caselli, R.J., Kukull, Walter A., McKeel, D., Morris, J.C., Hulette, C.M., Schmechel, D., Reiman, E.M., Rogers, J. & Stephan, D.A. (**2008**) Altered neuronal gene expression in brain regions differentially affected by Alzheimer's disease: a reference data set. *Physiological Genomics*, 33(2), 240–256.

Lin, H.-C., Hsieh, H.-M., Chen, Y.-H. & Hu, M.-L. (**2009**) S-Adenosylhomocysteine increases [beta]-amyloid formation in BV-2 microglial cells by increased expressions of [beta]-amyloid precursor protein and presenilin 1 and by hypomethylation of these gene promoters. *NeuroToxicology*, 30(4), 622–627.

Lindell, S.G., Yuan, Q., Zhou, Z., Goldman, D., Thompson, R.C., Lopez, J.F., Suomi, S.J., Higley, J.D. & Barr, C.S. (**2012**) The Serotonin Transporter Gene Is a Substrate for Age and Stress Dependent Epigenetic Regulation in Rhesus Macaque Brain: Potential Roles in Genetic Selection and Gene × Environment Interactions. *Development and Psychopathology*, 24(Special Issue 04), 1391–1400.

Lister, R., Pelizzola, M., Dowen, R.H., Hawkins, R.D., Hon, G., Tonti-Filippini, J., Nery, J.R., Lee, L., Ye, Z., Ngo, Q.-M., Edsall, L., Antosiewicz-Bourget, J., Stewart, R., Ruotti, V., Millar, A.H., Thomson, J.A., Ren, B. & Ecker, J.R. (**2009**) Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature*, 462(7271), 315–322.

Liu, Y., Aryee, M.J., Padyukov, L., Fallin, M.D., Hesselberg, E., Runarsson, A., Reinius, L., Acevedo, N., Taub, M., Ronninger, M., Shchetynsky, K., Scheynius, A., Kere, J., Alfredsson, L., Klareskog, L., Ekström, T.J. & Feinberg, A.P. (**2013**) Epigenome-wide association data implicate DNA methylation as an intermediary of genetic risk in rheumatoid arthritis. *Nature Biotechnology*, 31(2), 142–147.

Long, J.M. & Lahiri, D.K. (**2011**) MicroRNA-101 downregulates Alzheimer's amyloid-? precursor protein levels in human cell cultures and is differentially expressed. *Biochemical and Biophysical Research Communications*, 404(4), 889–895.

Lorincz, M.C., Dickerson, D.R., Schmitt, M. & Groudine, M. (**2004**) Intragenic DNA methylation alters chromatin structure and elongation efficiency in mammalian cells. *Nat Struct Mol Biol*, 11(11), 1068–1075.

Loring, J.F., Wen, X., Lee, J.M., Seilhamer, J. & Somogyi, R. (**2001**) A Gene Expression Profile of Alzheimer's Disease. *DNA and Cell Biology*, 20(11), 683–695.

Lotrich, F.E. & Pollock, B.G. (**2004**) Meta-analysis of serotonin transporter polymorphisms and affective disorders. *Psychiatric Genetics*, 14(3), 121–129.

Lotrich, F.E., Pollock, B.G., Kirshner, M., Ferrell, R.F. & Reynolds III, C.F. (**2008**) Serotonin transporter genotype interacts with paroxetine plasma levels to influence depression treatment response in geriatric patients. *Journal of Psychiatry & Neuroscience : JPN*, 33(2), 123–130.

Luchsinger, J.A., & Mayeux, R. (**2004**) Dietary factors and Alzheimer's disease. *The Lancet Neurology*, 3(10), 579–587.

Luger, K., Mäder, A.W., Richmond, R.K., Sargent, D.F. & Richmond, T.J. (**1997**) Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature*, 389(6648), 251–260.

Lumey, L.H., Terry, M.B., Delgado-Cruzata, L., Liao, Y., Wang, Q., Susser, E., McKeague, I. & Santella, R.M. (**2012**) Adult Global DNA Methylation in Relation to Pre-Natal Nutrition. *International Journal of Epidemiology*, 41(1), 116–123.

Lv, H., Jia, L. & Jia, J. (**2008**) Promoter polymorphisms which modulate APP expression may increase susceptibility to Alzheimer's disease. *Neurobiology of Aging*, 29(2), 194–202.

Malmgren, H., Steén-Bondeson, M.L., Gustavson, K.H., Seémanova, E., Holmgren, G., Oberlé, I., Mandel, J.L., Pettersson, U. & Dahl, N. (**1992**) Methylation and mutation patterns in the fragile X syndrome. *American Journal of Medical Genetics*, 43(1-2), 268–278.

Maloney, B., Ge, Y.-W., Greig, N. & Lahiri, D.K. (**2004**) Presence of a "CAGA box" in the APP gene unique to amyloid plaque-forming species and absent in all APLP-1/2 genes: implications in Alzheimer's disease. *The FASEB Journal*.

Mani, S.T. & Thakur, M.K. (**2006**) In the cerebral cortex of female and male mice, amyloid precursor protein (APP) promoter methylation is higher in females and differentially regulated by sex steroids. *Brain Research*, 1067(1), 43–47.

Mastroeni, D., Grover, A., Delvaux, E., Whiteside, C., Coleman, P.D. & Rogers, J. (**2010**) Epigenetic changes in Alzheimer's disease: Decrements in DNA methylation. *Neurobiology of Aging*, Articel in Press [Epub ahead of print](31), 2025–2037.

Mastroeni, D., Grover, A., Delvaux, E., Whiteside, C., Coleman, P.D. & Rogers, J. (**2011**) Epigenetic mechanisms in Alzheimer's disease. *Neurobiology of Aging*, 32(7), 1161–1180.

Mastroeni, D., McKee, A., Grover, A., Rogers, J. & Coleman, P.D. (**2009**) Epigenetic Differences in Cortical Neurons from a Pair of Monozygotic Twins Discordant for Alzheimer's Disease. *PLoS ONE*, 4(8), e6617.

Matsui, T., Ingelsson, M., Fukumoto, H., Ramasamy, K., Kowa, H., Frosch, M.P., Irizarry, M.C. & Hyman, B.T. (**2007**) Expression of APP pathway mRNAs and proteins in Alzheimer's disease. *Brain Research*, 1161, 116–123.

Mattson, M.P. (**1997**) Cellular actions of beta-amyloid precursor protein and its soluble and fibrillogenic derivatives. *Physiological Reviews*, 77(4), 1081–1132.

Maunakea, A.K., Nagarajan, R.P., Bilenky, M., Ballinger, T.J., D/'Souza, C., Fouse, S.D., Johnson, B.E., Hong, C., Nielsen, C., Zhao, Y., Turecki, G., Delaney, A., Varhol, R., Thiessen, N., Shchors, K., Heine, V.M., Rowitch, D.H., Xing, X., Fiore, C., Schillebeeckx, M., Jones, S.J.M., Haussler, D., Marra, M.A., Hirst, M., Wang, T. & Costello, J.F. (**2010**) Conserved role of intragenic DNA methylation in regulating alternative promoters. *Nature*, 466(7303), 253–257.

Mayeux, R. (**2003**) Epidemiology of Neurodegeneration. *Annual Review of Neuroscience*, 26(1), 81–104.

McCarthy, M.M., Auger, A.P., Bale, T.L., De Vries, G.J., Dunn, G.A., Forger, N.G., Murray, E.K., Nugent, B.M., Schwarz, J.M. & Wilson, M.E. (**2009**) The Epigenetics of Sex Differences in the Brain. *J. Neurosci.*, 29(41), 12815–12823.

McFarlane, A., Clark, C.R., Bryant, R.A., Williams, L.M., Niaura, R., Paul, R.H., Hitsman, B.L., Stroud, L., Alexander, D.M. & Gordon, E. (**2005**) The impact of early life stress on psychophysiological, personality and behavioral measures in 740 non-clinical subjects. *Journal of Integrative Neuroscience*, 04(01), 27–40.

McGowan, P.O., Sasaki, A., Huang, T.C.T., Unterberger, A., Suderman, M., Ernst, C., Meaney, M.J., Turecki, G. & Szyf, M. (**2008**) Promoter-Wide Hypermethylation of the Ribosomal RNA Gene Promoter in the Suicide Brain. *PLoS ONE*, 3(5), e2085.

McIlroy, S. & Craig, D. (**2004**) Neurobiology and genetics of behavioural syndromes of Alzheimer's disease. *Current Alzheimer Research*, 1(2), 135–142.

Meaney, M.J. (**2010**) Epigenetics and the Biological Definition of Gene × Environment Interactions. *Child Development*, 81(1), 41–79.

Mega, M.S., Cummings, J.L., Fiorello, T. & Gornbein, J. (**1996**) The spectrum of behavioral changes in Alzheimer's disease. *Neurology*, 46(1), 130–135.

Meissner, A., Mikkelsen, T.S., Gu, H., Wernig, M., Hanna, J., Sivachenko, A., Zhang, Xiaolan, Bernstein, B.E., Nusbaum, C., Jaffe, D.B., Gnirke, A., Jaenisch, R. & Lander, E.S. (**2008**) Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature*, 454(7205), 766–770.

Menéndez-González, M., Pérez-Pinera, P., Martínez-River, M., Calatayud, M.T. & Blázquez Menes, B. (**2005**) APP Processing and the APP-KPI Domain Involvement in the Amyloid Cascade. *Neurodegenerative Diseases*, 2(6), 277–283.

Micheli, D., Bonvicini, C., Rocchi, A., Ceravolo, R., Mancuso, M., Tognoni, G., Gennarelli, M., Siciliano, G. & Murri, L. (**2006**) No evidence for allelic association of serotonin 2A receptor and transporter gene polymorphisms with depression in Alzheimer disease. *Journal of Alzheimer's Disease*, 10(4), 371–378.

Milici, A., Salbaum, J.M. & Beyereuther, K. (**1990**) Study of the Alzheimer's A4 precursor gene promoter region by genomic sequencing using taq polymerase. *Biochemical and Biophysical Research Communications*, 169(1), 46–50.

Mill, J & Petronis, A (**2007**) Molecular studies of major depressive disorder: the epigenetic perspective. *Mol Psychiatry*, 12(9), 799–814.

Mill, J & Petronis, Arturas (**2008**) Pre- and peri-natal environmental risks for attention-deficit hyperactivity disorder (ADHD): the potential role of epigenetic processes in mediating susceptibility. *Journal of Child Psychology and Psychiatry*, 49(10), 1020–1030.

Mill, J., Tang, T., Kaminsky, Z., Khare, T., Yazdanpanah, S., Bouchard, L., Jia, P., Assadzadeh, A., Flanagan, J., Schumacher, A., Wang, S.-C. & Petronis, A. (**2008**) Epigenomic Profiling Reveals DNA-Methylation Changes Associated with Major Psychosis. *The American Journal of Human Genetics*, 82(3), 696–711.

Miller, C.A. & Sweatt, J.D. (**2007**) Covalent Modification of DNA Regulates Memory Formation. *Neuron*, 53(6), 857–869.

Miller, D.B. & O'Callaghan, J.P. (**2008**) Do early-life insults contribute to the late-life development of Parkinson and Alzheimer diseases? *Metabolism*, 57(Supplement 2), S44–S49.

Miller, J.A., Oldham, M.C. & Geschwind, D.H. (**2008**) A Systems Level Analysis of Transcriptional Changes in Alzheimer's Disease and Normal Aging. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 28(6), 1410–1420.

Miranda, T.B. & Jones, P.A. (**2007**) DNA methylation: The nuts and bolts of repression. *Journal of Cellular Physiology*, 213(2), 384–390.

Moceri, V.M., Kukull, W. A., Emanuel, I., van Belle, G. & Larson, E.B. (**2000**) Early-life risk factors and the development of Alzheimer's disease. *Neurology*, 54(2), 415.

Modrego, P.J. (**2010**) Depression in Alzheimer's Disease. Pathophysiology, Diagnosis, and Treatment. *Journal of Alzheimer's Disease*, 21(4), 1077–1087.

Mohandas, E., Rajmohan, V. & Raghunath, B. (**2009**) Neurobiology of Alzheimer's disease. *Indian Journal of Psychiatry*, 51(1), 55–61.

Moir, R.D., Lynch, T., Bush, A.I., Whyte, S., Henry, A., Portbury, S., Multhaup, G., Small, D.H., Tanzi, R.E., Beyreuther, K. & Masters, C.L. (**1998**) Relative Increase in Alzheimer's Disease of Soluble Forms of Cerebral Aβ Amyloid Protein Precursor Containing the Kunitz Protease Inhibitory Domain. *Journal of Biological Chemistry*, 273(9), 5013–5019.

Montine, T.J., Phelps, C.H., Beach, T.G., Bigio, E.H., Cairns, N.J., Dickson, D.W., Duyckaerts, C., Frosch, M.P., Masliah, E., Mirra, S.S., Nelson, P.T., Schneider, J.A., Thal, D.R., Trojanowski, J.Q., Vinters, H.V. & Hyman, B.T. (**2012**) National Institute on Aging-Alzheimer's Association guidelines for the neuropathologic assessment of Alzheimer's disease: a practical approach. *Acta Neuropathologica*, 123(1), 1–11.

Morgan, A.R., Hamilton, G., Turic, D., Jehu, L., Harold, D., Abraham, R., Hollingworth, P., Moskvina, V., Brayne, C., Rubinsztein, D.C., Lynch, A., Lawlor, B., Gill, M., O'Donovan, M., Powell, J., Lovestone, S., Williams, J. & Owen, M.J. (**2008**) Association analysis of 528 intra-genic SNPs in a region of chromosome 10 linked to late onset Alzheimer's disease. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, 147B(6), 727–731.

Mortensen, O.V., Thomassen, M., Larsen, M.B., Whittemore, S.R. & Wiborg, O. (**1999**) Functional analysis of a novel human serotonin transporter gene promoter in immortalized raphe cells. *Molecular Brain Research*, 68(1–2), 141–148.

Mortimer, J.A., Snowdon, D.A. & Markesbery, W.R. (**2003**) Head Circumference, Education and Risk of Dementia: Findings from the Nun Study. *Journal of Clinical and Experimental Neuropsychology*, 25(5), 671.

Mortimer, J.A., Snowdon, D.A. & Markesbery, W.R. (**2008**) Small Head Circumference is Associated With Less Education in Persons at Risk for Alzheimer Disease in Later Life. *Alzheimer Disease & Associated Disorders July/September 2008*, 22(3), 249–254.

Mossello, E., Boncinelli, M., Caleri, V., Cavallini, M.C., Palermo, E., Di Bari, M., Tilli, S., Sarcone, E., Simoni, D., Biagini, C.A., Masotti, G. & Marchionni, N. (**2008**) Is Antidepressant Treatment Associated with Reduced Cognitive Decline in Alzheimer's Disease? *Dementia and Geriatric Cognitive Disorders*, 25(4), 372–379.

Mowla, A., Mosavinasab, M., Haghshenas, H. & Haghighi, A.B. (**2007**) Does Serotonin Augmentation Have Any Effect on Cognition and Activities of Daily Living in Alzheimer's Dementia? *Journal of Clinical Psychopharmacology*, 27(5), 484–487.

Münzel, M., Globisch, D., Brückl, T., Wagner, M., Welzmiller, V., Michalakis, S., Müller, M., Biel, M. & Carell, T. (**2010**) Quantification of the Sixth DNA Base Hydroxymethylcytosine in the Brain. *Angewandte Chemie International Edition*, 49(31), 5375–5377.

Murgatroyd, C., Patchev, A.V., Wu, Y., Micale, V., Bockmuhl, Y., Fischer, D., Holsboer, F., Wotjak, C.T., Almeida, O.F.X. & Spengler, D. (**2009**) Dynamic DNA methylation programs persistent adverse effects of early-life stress. *Nat Neurosci*, 12(12), 1559–1566.

Nabel, C.S. & Kohli, R.M. (**2011**) Demystifying DNA Demethylation. *Science*, 333(6047), 1229–1230.

Naj, A.C., Jun, G., Beecham, G.W., Wang, L.-S., Vardarajan, B.N., et al. (**2011**) Common variants at MS4A4/MS4A6E, CD2AP, CD33 and EPHA1 are associated with late-onset Alzheimer's disease. *Nature Genetics*, 43(5), 436–441.

Nakamura, M., Ueno, S., Sano, A. & Tanabe, H. (**2000**) The human serotonin transporter gene linked polymorphism (5-HTTLPR) shows ten novel allelic variants. *Molecular psychiatry*, 5(1), 32–38.

Nan, X, Ng, H.H., Johnson, C.A., Laherty, C.D., Turner, B.M., Eisenman, R.N. & Bird, A (**1998**) Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature*, 393(6683), 386–389.

Narayan, P. & Dragunow, M. (**2010**) Pharmacology of epigenetics in brain disorders. *British Journal of Pharmacology*, 159(2), 285–303.

Nee, L.E. & Lippa, C.F. (**1999**) Alzheimer's Disease in 22 Twin Pairs 13-Year Follow-Up: Hormonal, Infectious and Traumatic Factors. *Dementia and Geriatric Cognitive Disorders*, 10(2), 148–151.

Nestor, C., Ruzov, A., Meehan, R. & Dunican, D. (**2010**) Enzymatic approaches and bisulfite sequencing cannot distinguish between 5-methylcytosine and 5-hydroxymethylcytosine in DNA. *BioTechniques*, 48(4), 317–319.

Nestor, C.E., Ottaviano, R., Reddington, J., Sproul, D., Reinhardt, D., Dunican, D., Katz, E., Dixon, J.M., Harrison, D.J. & Meehan, R.R. (**2011**) Tissue type is a major modifier of the 5-hydroxymethylcytosine content of human genes. *Genome Research*.

Neumann, H. & Daly, M.J. (**2013**) Variant TREM2 as Risk Factor for Alzheimer's Disease. *New England Journal of Medicine*, 368(2), 182–184.

Niculescu, M.D., Craciunescu, C.N. & Zeisel, S.H. (**2006**) Dietary choline deficiency alters global and gene-specific DNA methylation in the developing hippocampus of mouse fetal brains. *The FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, 20(1), 43–49.

Numata, S., Ye, T., Hyde, T.M., Guitart-Navarro, X., Tao, R., Wininger, M., Colantuoni, C., Weinberger, D.R., Kleinman, J.E. & Lipska, B.K. (**2012**) DNA Methylation Signatures in Development and Aging of the Human Prefrontal Cortex. *The American Journal of Human Genetics*, 90(2), 260–272.

Ohlsson, R., Renkawitz, R. & Lobanenkov, V. (**2001**) CTCF is a uniquely versatile transcription regulator linked to epigenetics and disease. *Trends in Genetics*, 17(9), 520–527.

Oliveira, J.R.M., Shimokomaki, C.M., Brito-Marques, P.R., Gallindo, R.M., Okuma, M., Maia, L.G.S., Otto, P.A., Passos-Bueno, M.R. & Zatz, M. (**1999**) The association of the short variant of the 5-HTTPLR polymorphism and the apoE ε4 allele does not increase the risk for late onset Alzheimer's disease. *Molecular Psychiatry*, 4(1), 19–20.

Olsson, C.A., Foley, D.L., Parkinson-Bates, M., Byrnes, G., McKenzie, M., Patton, G.C., Morley, R., Anney, R.J.L., Craig, J.M. & Saffery, R. (**2010**) Prospects for epigenetic research within cohort studies of psychological disorder: a pilot investigation of a peripheral cell marker of epigenetic risk for depression. *Biological psychology*, 83(2), 159–165.

Ouchi, Y., Yoshikawa, E., Futatsubashi, M., Yagi, S., Ueki, T. & Nakamura, K. (**2009**) Altered Brain Serotonin Transporter and Associated Glucose Metabolism in Alzheimer Disease. *Journal of Nuclear Medicine*, 50(8), 1260–1266.

Oyama, F., Cairns, N.J., Shimada, H., Oyama, R., Titani, K. & Ihara, Y. (**1994**) Down's Syndrome: Up-Regulation of β-Amyloid Protein Precursor and τ mRNAs and Their Defective Coordination. *Journal of Neurochemistry*, 62(3), 1062–1066.

Parsey, R.V., Hastings, R.S., Oquendo, M.A., Hu, X., Goldman, D., Huang, Yung-yu, Simpson, N., Arcement, J., Huang, Yiyun, Ogden, R.T., Van Heertum, R.L., Arango, V. & Mann, J.J. (**2006**) Effect of a Triallelic Functional Polymorphism of the Serotonin-Transporter-Linked Promoter Region on Expression of Serotonin Transporter in the Human Brain. *American Journal of Psychiatry*, 163(1), 48–51.

Pastor, W.A., Pape, U.J., Huang, Yun, Henderson, H.R., Lister, R., Ko, M., McLoughlin, E.M., Brudno, Y., Mahapatra, S., Kapranov, P., Tahiliani, M., Daley, G.Q., Liu, X.S., Ecker, J.R., Milos, P.M., Agarwal, S. & Rao, A. (**2011**) Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. *Nature*, 473(7347), 394–397.

Paulsen, M. (**2008**) DNA Methylation & the Mammalian Genome. In *Epigenetics*. Editor: Tost, J., Horizon Scientific Press.

Perry, V.H., Cunningham, C. & Holmes, Clive (**2007**) Systemic infections and inflammation affect chronic neurodegeneration. *Nature Reviews Immunology*, 7(2), 161–167.

Petronis, A. (**2010**) Epigenetics as a unifying principle in the aetiology of complex traits and diseases. *Nature*, 465(7299), 721–727.

Petronis, A. (**2001**) Human morbid genetics revisited: relevance of epigenetics. *Trends in Genetics*, 17(3), 142–146.

Philibert, R.A., Sandhu, H., Hollenbeck, N., Gunter, T., Adams, W. & Madan, A. (**2008**) The relationship of 5HTT (SLC6A4) methylation and genotype on mRNA expression and liability to major depression and alcohol dependence in subjects from the Iowa Adoption Studies. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, 147B(5), 543–549.

Piaceri, I., Nacmias, B. & Sorbi, S. (**2013**) Genetics of familial and sporadic Alzheimer's disease. *Frontiers in Bioscience (Elite Edition)*, 5, 167–177.

Pidsley, R. & Mill, J. (**2011**) Epigenetic studies of psychosis: current findings, methodological approaches, and implications for postmortem research. *Biological Psychiatry*, 69(2), 146–156.

Podlisny, M.B., Tolan, D.R. & Selkoe, D. J. (**1991**) Homology of the amyloid beta protein precursor in monkey and human supports a primate model for beta amyloidosis in Alzheimer's disease. *The American Journal of Pathology*, 138(6), 1423–1435.

Pogribny, I.P. & Vanyushin, B.F. (**2010**) Age-Related Genomic Hypomethylation. In T. O. Tollefsbol, ed. *Epigenetics of Aging*. Springer New York, pp. 11–27.

Politis, A., Olgiati, P., Malitas, P., Albani, D., Signorini, A., Polito, L., De Mauro, S., Zisaki, A., Piperi, C., Stamouli, E., Mailis, A., Batelli, S., Forloni, G., De Ronchi, D., Kalofoutis, A., Liappas, I. & Serretti, A. (**2010**) Vitamin B12 Levels in Alzheimer's Disease: Association with Clinical Features and Cytokine Production. *Journal of Alzheimer's Disease*, 19(2), 481–488.

Pollwein, P. (**1993**) Overlapping Binding Sites of Two Different Transcription Factors in the Promoter of the Human Gene for the Alzheimer Amyloid Precursor Protein. *Biochemical and Biophysical Research Communications*, 190(2), 637–647.

Pritchard, A.L., Pritchard, C.W., Bentham, P. & Lendon, C.L. (**2007**) Role of Serotonin Transporter Polymorphisms in the Behavioural and Psychological Symptoms in Probable Alzheimer Disease Patients. *Dementia and Geriatric Cognitive Disorders*, 24(3), 201–206.

Profenno, L.A., Porsteinsson, A.P. & Faraone, S.V. (**2010**) Meta-Analysis of Alzheimer's Disease Risk with Obesity, Diabetes, and Related Disorders. *Biological Psychiatry*, 67(6), 505–512.

Quaranta, D., Bizzarro, A., Marra, C., Vita, M.G., Seripa, D., Pilotto, A., Sebastiani, V., Mecocci, P. & Masullo, C. (**2009**) Psychotic Symptoms in Alzheimer's Disease and 5-HTTLPR Polymorphism of the Serotonin Transporter Gene: Evidence for an Association. *Journal of Alzheimer's Disease*, 16(1), 173–180.

Querfurth, H.W., Jiang, J., Xia, W. & Selkoe, D.J. (**1999**) Enhancer function and novel DNA binding protein activity in the near upstream βAPP gene promoter. *Gene*, 232(1), 125–141.

Querfurth, H.W. & LaFerla, F.M. (**2010**) Alzheimer's Disease. *N Engl J Med*, 362(4), 329–344.

Quitschke, W.W & Goldgaber, D. (**1992**) The amyloid beta-protein precursor promoter. A region essential for transcriptional activity contains a nuclear factor binding domain. *The Journal of Biological Chemistry*, 267(24), 17362–17368.

Rademakers, R., Cruts, M. & van Broeckhoven, C. (**2004**) The role of tau (MAPT) in frontotemporal dementia and related tauopathies. *Human Mutation*, 24(4), 277–295.

Rakyan, V.K., Down, T.A., Balding, D.J. & Beck, S. (**2011**) Epigenome-wide association studies for common human diseases. *Nat Rev Genet*, 12(8), 529–541.

Ramsahoye, B.H., Biniszkiewicz, D., Lyko, F., Clark, V., Bird, A.P. & Jaenisch, R. (**2000**) Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. *Proceedings of the National Academy of Sciences*, 97(10), 5237 –5242.

Rao, J.S., Keleshian, V.L., Klein, S. & Rapoport, S.I. (**2012**) Epigenetic modifications in frontal cortex from Alzheimer's disease and bipolar disorder patients. *Translational Psychiatry*, 2(7), e132.

Rath, P.C. & Kanungo, M.S. (**1989**) Methylation of repetitive DNA sequences in the brain during aging of the rat. *FEBS Letters*, 244(1), 193–198.

Rausch, J.L., Johnson, M.E., Fei, Y.-J., Li, J.Q., Shendarkar, N., Hobby, H.M., Ganapathy, V. & Leibach, F.H. (**2002**) Initial conditions of serotonin transporter kinetics and genotype: influence on SSRI treatment trial outcome. *Biological Psychiatry*, 51(9), 723–732.

Ravaglia, G., Forti, P., Maioli, F., Martelli, M., Servadei, L., Brunetti, N., Porcellini, E. & Licastro, F. (**2005**) Homocysteine and folate as risk factors for dementia and Alzheimer disease. *The American Journal of Clinical Nutrition*, 82(3), 636–643.

Ressler, K.J. & Nemeroff, C.B. (**1999**) Role of norepinephrine in the pathophysiology and treatment of mood disorders. *Biological Psychiatry*, 46(9), 1219–1233.

Rizwana, R. & Hahn, P.J. (**1999**) CpG methylation reduces genomic instability. *Journal of Cell Science*, 112(24), 4513–4519.

Robertson, A.B., Dahl, J.A., Vågbø, C.B., Tripathi, P., Krokan, H.E. & Klungland, A. (**2011**) A novel method for the efficient and selective identification of 5-hydroxymethylcytosine in genomic DNA. *Nucleic Acids Research*.

Rogaev, E.I., Lukiw, W.J., Lavrushina, O., Rogaeva, E.A. & St. George-Hyslop, P.H. (**1994**) The Upstream Promoter of the [beta]-Amyloid Precursor Protein Gene (APP) Shows Differential Patterns of Methylation in Human Brain. *Genomics*, 22(2), 340–347.

Rollins, R.A., Haghighi, Fatemeh, Edwards, J.R., Das, R., Zhang, M.Q., Ju, J. & Bestor, T.H. (**2006**) Large-scale structure of genomic methylation patterns. *Genome Research*, 16(2), 157–163.

Roman, D.L., Walline, C.C., Rodriguez, G.J. & Barker, E.L. (**2003**) Interactions of antidepressants with the serotonin transporter: a contemporary molecular analysis. *European Journal of Pharmacology*, 479(1–3), 53–63.

Rovelet-Lecrux, A., Frebourg, T., Tuominen, H., Majamaa, K., Campion, D. & Remes, A.M. (**2007**) APP locus duplication in a Finnish family with dementia and intracerebral haemorrhage. *Journal of Neurology, Neurosurgery & Psychiatry*, 78(10), 1158–1159.

Rovelet-Lecrux, A., Hannequin, D., Raux, G., Meur, N.L., Laquerrière, A., Vital, A., Dumanchin, C., Feuillette, S., Brice, A., Vercelletto, M., Dubas, F., Frebourg, T. & Campion, D. (**2006**) APP locus duplication causes autosomal dominant early-onset Alzheimer disease with cerebral amyloid angiopathy. *Nature Genetics*, 38(1), 24–26.

Rumble, B., Retallack, R., Hilbich, C., Simms, G., Multhaup, G., Martins, R., Hockey, A., Montgomery, P., Beyreuther, K. & Masters, C.L. (**1989**) Amyloid A4 Protein and Its Precursor in Down's Syndrome and Alzheimer's Disease. *New England Journal of Medicine*, 320(22), 1446–1452.

Sakamoto, K., Karelina, K. & Obrietan, K. (**2011**) CREB: a multifaceted regulator of neuronal plasticity and protection. *Journal of Neurochemistry*, 116(1), 1–9.

Sandbrink, R., Banati, R., Masters, C.L., Beyreuther, K. & König, G. (**1993**) Expression of L-APP mRNA in Brain Cells. *Annals of the New York Academy of Sciences*, 695(1), 183–189.

Sandbrink, R., Monning, U., Masters, C.L. & Beyreuther, K. (**1997**) Expression of the APP Gene Family in Brain Cells, Brain Development and Aging. *Gerontology*, 43(1-2), 119–131.

Saunders, A.M, Strittmatter, W.J., Schmechel, D, George-Hyslop, P.H., Pericak-Vance, M.A., Joo, S.H., Rosi, B.L., Gusella, J.F., Crapper-MacLachlan, D.R. & Alberts, M.J. (**1993**) Association of apolipoprotein E allele epsilon 4 with late-onset familial and sporadic Alzheimer's disease. *Neurology*, 43(8), 1467–1472.

Saxonov, S., Berg, P. & Brutlag, D.L. (**2006**) A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. *Proceedings of the National Academy of Sciences of the United States of America*, 103(5), 1412–1417.

Scarmeas, N., Luchsinger, Jose A, Schupf, N., Brickman, A.M., Cosentino, S., Tang, M.X. & Stern, Yaakov (**2009**) Physical activity, diet, and risk of Alzheimer disease. *JAMA: The Journal of the American Medical Association*, 302(6), 627–637.

Schjeide, B.-M.M., McQueen, M.B., Mullin, K., DiVito, J., Hogan, M.F., Parkinson, M., Hooli, B., Lange, C., Blacker, D., Tanzi, R.E. & Bertram, L. (**2009**) Assessment of Alzheimer's disease case–control associations using family-based methods. *neurogenetics*, 10(1), 19–25.

Schofield, P.W., Logroscino, G., Andrews, H.F., Albert, S. & Stern, Y. (**1997**) An association between head circumference and Alzheimer's disease in a population-based study of aging and dementia. *Neurology*, 49(1), 30–37.

Schulz, R.-J. (**2007**) Homocysteine as a biomarker for cognitive dysfunction in the elderly. *Current Opinion in Clinical Nutrition and Metabolic Care*, 10(6), 718–723.

Schumacher (**2011**) Aging Epigenetis. In *Handbook of Epigenetics: The New Molecular and Medical Genetics,* 2011 Edited by: Trygve Tollefsbol, Elsevier Inc.

Schwob, N.G., Nalbantoglu, J., Hastings, K.E.M., Mikkelsen, T. & Cashman, N.R. (**1990**) *DNA cytosine methylation in brain of patients with Alzheimer's disease. *Annals of Neurology*, 28(1), 91–94.

Selkoe, D.J. (**2002**) Alzheimer's Disease Is a Synaptic Failure. *Science*, 298(5594), 789–791.

Seripa, D., Franceschi, M., D'Onofrio, G., Panza, F., Cascavilla, L., Paris, F., Placentino, G., Matera, M.G., Solfrizzi, V. & Pilotto, A. (**2008**) Polymorphism C in the serotonin transporter gene (SLC6A4) in questionable dementia and Alzheimer's disease. *Neuroscience Letters*, 438(3), 335–339.

Seshadri, S., Beiser, A., Selhub, J., Jacques, P.F., Rosenberg, I.H., D'Agostino, R.B., Wilson, P.W.F. & Wolf, P.A. (**2002**) Plasma Homocysteine as a Risk Factor for Dementia and Alzheimer's Disease. *New England Journal of Medicine*, 346(7), 476–483.

Shea, T.B. & Rogers, E. (**2002**) Folate quenches oxidative damage in brains of apolipoprotein E-deficient mice: augmentation by vitamin E. *Brain Research. Molecular Brain Research*, 108(1-2), 1–6.

Shenker, N. & Flanagan, J.M. (**2012**) Intragenic DNA methylation: implications of this epigenetic mechanism for cancer research. *British Journal of Cancer*, 106(2), 248–253.

Sherrington, R., Rogaev, E.I., Liang, Y., Rogaeva, E.A., Levesque, G., Ikeda, M., Chi, H., Lin, C., Li, G., Holman, K., Tsuda, T., Mar, L., Foncin, J.-F., Bruni, A.C., Montesi, M.P., Sorbi, S., Rainero, I., Pinessi, L., Nee, L., Chumakov, I., Pollen, D., Brookes, A., Sanseau, P., Polinsky, R.J., Wasco, W., Da Silva, H.A.R., Haines, J.L., Pericak-Vance, M.A., Tanzi, R. E., Roses, A.D., Fraser, P.E., Rommens, J.M. & St George-Hyslop, P.H. (**1995**) Cloning of a gene bearing missense mutations in early-onset familial Alzheimer's disease. *Nature*, 375(6534), 754–760.

Shioe, K., Ichimiya, T., Suhara, T., Takano, A., Sudo, Y., Yasuno, F., Hirano, M., Shinohara, M., Kagami, M., Okubo, Y., Nankai, M. & Kanba, S. (**2003**) No association between genotype of the promoter region of serotonin transporter gene and serotonin transporter binding in human brain measured by PET. *Synapse*, 48(4), 184–188.

Shukla, S., Kavak, E., Gregory, M., Imashimizu, M., Shutinoski, B., Kashlev, M., Oberdoerffer, P., Sandberg, R. & Oberdoerffer, S. (**2011**) CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing. *Nature*, 479(7371), 74–79.

Siegfried, Z., Eden, S., Mendelsohn, M., Feng, X., Tsuberi, B.-Z. & Cedar, H. (**1999**) DNA methylation represses transcription in vivo. *Nature Genetics*, 22(2), 203–206.

Siegmund, K.D., Connor, C.M., Campan, M., Long, T.I., Weisenberger, D.J., Biniszkiewicz, D., Jaenisch, R., Laird, P.W. & Akbarian, S. (**2007**) DNA Methylation in the Human Cerebral Cortex Is Dynamically Regulated throughout the Life Span and Involves Differentiated Neurons. *PLoS ONE*, 2(9), e895.

Silva, P.N.O., Gigek, C.O., Leal, M.F., Bertolucci, P.H.F., de Labio, R.W., Payão, S.L.M. & Smith, M. de A.C. (**2008**) Promoter Methylation Analysis of SIRT3, SMARCA5, HTERT and CDH1 Genes in Aging and Alzheimer's Disease. *Journal of Alzheimer's Disease*, 13(2), 173–176.

Sleegers, K., Brouwers, N., Gijselinck, I., Theuns, J., Goossens, D., Wauters, J., Del-Favero, J., Cruts, Marc, Duijn, C.M. van & Broeckhoven, C.V. (**2006**) APP duplication is sufficient to cause early onset Alzheimer's dementia with cerebral amyloid angiopathy. *Brain*, 129(11), 2977 –2983.

Song, C.-X., Sun, Yao, Dai, Q., Lu, X., Yu, M., Yang, C. & He, C. (**2011a**) Detection of 5-Hydroxymethylcytosine in DNA by Transferring a Keto-Glucose by Using T4 Phage β-Glucosyltransferase. *ChemBioChem*, 12(11), 1682–1685.

Song, C.-X., Szulwach, K.E., Fu, Y., Dai, Q., Yi, C., Li, Xuekun, Li, Y., Chen, C.-H., Zhang, W., Jian, X., Wang, Jing, Zhang, Li, Looney, T.J., Zhang, B., Godley, L.A., Hicks, L.M., Lahn, B.T., Jin, P. & He, C. (**2010**) Selective chemical labeling reveals the genome-wide distribution of 5-hydroxymethylcytosine. *Nat Biotech*, advance online publication.

Song, C.-X., Yu, M., Dai, Q. & He, C. (**2011b**) Detection of 5-hydroxymethylcytosine in a combined glycosylation restriction analysis (CGRA) using restriction enzyme Taq[alpha]I. *Bioorganic & Medicinal Chemistry Letters*, 21(17), 5075–5077.

Starkstein, S.E., Mizrahi, R. & Power, B.D. (**2008**) Depression in Alzheimer's disease: Phenomenology, clinical correlates and treatment. *International Review of Psychiatry*, 20(4), 382–388.

Stewart, W.F., Schwartz, B.S., Davatzikos, C., Shen, D., Liu, D., Wu, X, Todd, A C, Shi, W., Bassett, S. & Youssem, D. (**2006**) Past adult lead exposure is linked to neurodegeneration measured by brain MRI. *Neurology*, 66(10), 1476–1484.

Stewart, W.F., Schwartz, B.S., Simon, D., Kelsey, K. & Todd, A.C. (**2002**) ApoE genotype, past adult lead exposure, and neurobehavioral function. *Environmental Health Perspectives*, 110(5), 501–505.

Sukonick, D.L., Pollock, B.G., Sweet, R.A., Mulsant, B.H., Rosen, J., Klunk, W.E., Kastango, K.B., DeKosky, S.T. & Ferrell, R.E. (**2001**) The 5-HTTPR*S/*L polymorphism and aggressive behavior in Alzheimer disease. *Archives of neurology*, 58(9), 1425–1428.

Sultana, R. & Butterfield, D.A. (**2010**) Role of Oxidative Stress in the Progression of Alzheimer's Disease. *Journal of Alzheimer's Disease*, 19(1), 341–353.

Sved, J. & Bird, A. (**1990**) The expected equilibrium of the CpG dinucleotide in vertebrate genomes under a mutation model. *Proceedings of the National Academy of Sciences*, 87(12), 4692–4696.

Sweet, R.A., Pollock, B.G., Sukonick, D.L., Mulsant, B.H., Rosen, J., Klunk, W.E., Kastango, K.B., DeKosky, S.T. & Ferrell, R.E. (**2001**) The 5-HTTPR Polymorphism Confers Liability to a Combined Phenotype of Psychotic and Aggressive Behavior in Alzheimer Disease. *International Psychogeriatrics*, 13(04), 401–409.

Swerdlow, R.H., Burns, J.M. & Khan, S.M. (**2010**) The Alzheimer's Disease Mitochondrial Cascade Hypothesis. *Journal of Alzheimer's Disease*.

Szulwach, K.E., Li, Xuekun, Li, Y., Song, C.-X., Wu, H., Dai, Q., Irier, H., Upadhyay, A.K., Gearing, M., Levey, A.I., Vasanthakumar, A., Godley, L.A., Chang, Q., Cheng, X., He, C. & Jin, P. (**2011**) 5-hmC-mediated epigenetic dynamics during postnatal neurodevelopment and aging. *Nat Neurosci*, 14(12), 1607–1616.

Tahiliani, M., Koh, K.P., Shen, Yinghua, Pastor, W.A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L.M., Liu, D.R., Aravind, L. & Rao, A. (**2009**) Conversion of 5-Methylcytosine to 5-Hydroxymethylcytosine in Mammalian DNA by MLL Partner TET1. *Science*, 324(5929), 930–935.

Talens, R.P., Boomsma, D.I., Tobi, E.W., Kremer, D., Jukema, J.W., Willemsen, G., Putter, H., Slagboom, P.E. & Heijmans, B.T. (**2010**) Variation, patterns, and temporal stability of DNA methylation: considerations for epigenetic epidemiology. *The FASEB Journal*, 24(9), 3135–3144.

Tejani-Butt, S.M., Yang, J. & Pawlyk, A.C. (**1995**) Altered serotonin transporter sites in Alzheimer's disease raphe and hippocampus. *NeuroReport*, 6(8), 1207–1210.

Tewhey, R., Warner, J.B., Nakano, M., Libby, B., Medkova, M., David, P.H., Kotsopoulos, S.K., Samuels, M.L., Hutchison, J.B., Larson, J.W., Topol, E.J., Weiner, M.P., Harismendy, O., Olson, J., Link, D.R. & Frazer, K.A. (**2009**) Microdroplet-based PCR enrichment for large-scale targeted sequencing. *Nature Biotechnology*, 27(11), 1025–1031.

The ENCODE Project (**2004**) The ENCODE (ENCyclopedia Of DNA Elements) Project. *Science*, 306(5696), 636–640.

The ENCODE Project (**2011**) A User's Guide to the Encyclopedia of DNA Elements (ENCODE). *PLoS Biology*, 9(4).

The ENCODE Project (**2007**) Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature*, 447(7146), 799–816.

Theuns, J. & Broeckhoven, C.V. (**2000**) Transcriptional regulation of Alzheimer's disease genes: implications for susceptibility. *Human Molecular Genetics*, 9(16), 2383–2394.

Theuns, J., Brouwers, N., Engelborghs, S., Sleegers, K., Bogaerts, V., Corsmit, E., De Pooter, T., van Duijn, C.M., De Deyn, P.P. & Van Broeckhoven, C. (**2006**) Promoter Mutations That Increase Amyloid Precursor-Protein Expression Are Associated with Alzheimer Disease. *The American Journal of Human Genetics*, 78(6), 936–946.

Thomas, A.J., Hendriksen, M., Piggott, M., Ferrier, I.N., Perry, E., Ince, P. & O'Brien, J.T. (**2006**) A study of the serotonin transporter in the prefrontal cortex in late-life depression and Alzheimer's disease with and without depression. *Neuropathology and Applied Neurobiology*, 32(3), 296–303.

Thurman, R.E., Rynes, E., Humbert, R., Vierstra, J., Maurano, M.T., et al. (**2012**) The accessible chromatin landscape of the human genome. *Nature*, 489(7414), 75–82.

Tohgi, H., Utsugisawa, K., Nagane, Y., Yoshimura, M., Genda, Y. & Ukitsu, M. (**1999a**) Reduction with age in methylcytosine in the promoter region -224~-101 of the amyloid precursor protein gene in autopsy human cortex. *Molecular Brain Research*, 70(2), 288–292.

Tohgi, H., Utsugisawa, K., Nagane, Y., Yoshimura, M., Ukitsu, M. & Genda, Y. (**1999b**) The methylation status of cytosines in a τ gene promoter region alters with age to downregulate transcriptional activity in human cerebral cortex. *Neuroscience Letters*, 275(2), 89–92.

Tra, J., Kondo, T., Lu, Q., Kuick, R., Hanash, S. & Richardson, B. (**2002**) Infrequent occurrence of age-dependent changes in CpG island methylation as detected by restriction landmark genome scanning. *Mechanisms of Ageing and Development*, 123(11), 1487–1503.

Trasler, J., Deng, Liyuan, Melnyk, S., Pogribny, I., Hiou-Tim, F., Sibani, S., Oakes, C., Li, E., James, S.J. & Rozen, R. (**2003**) Impact of Dnmt1 deficiency, with and without low folate diets, on tumor numbers and DNA methylation in Min mice. *Carcinogenesis*, 24(1), 39–45.

Truchot, L., Costes, S.N., Zimmer, L., Laurent, B., Bars, D.L., Thomas-Antérion, C., Croisile, B., Mercier, B., Hermier, M., Vighetto, A. & Krolak-Salmon, P. (**2007**) Up-regulation of hippocampal serotonin metabolism in mild cognitive impairment. *Neurology*, 69(10), 1012–1017.

Tsai, M.-C., Manor, O., Wan, Yue, Mosammaparast, N., Wang, J.K., Lan, F., Shi, Y., Segal, E. & Chang, H.Y. (**2010**) Long Noncoding RNA as Modular Scaffold of Histone Modification Complexes. *Science*, 329(5992), 689–693.

Tsang, S.W.Y., Lai, M.K.P., Francis, P.T., Wong, P.T.-H., Spence, I., Esiri, M.M., Keene, J., Hope, T. & Chen, C.P.L.-H. (**2003**) Serotonin transporters are preserved in the neocortex of anxious Alzheimer's disease patients. *NeuroReport*, 14(10), 1297–1300.

Tsukada, Y.-I. (**2012**) Hydroxylation Mediates Chromatin Demethylation. *Journal of Biochemistry*, 151(3), 229–246.

Tusnády, G.E., Simon, I., Váradi, A. & Arányi, T. (**2005**) BiSearch: primer-design and search tool for PCR on bisulfite-treated genomes. *Nucleic Acids Research*, 33(1), e9–e9.

Ueki, A., Ueno, H., Sato, N., Shinjo, H. & Morita, Y. (**2007**) Serotonin Transporter Gene Polymorphism and BPSD in Mild Alzheimer's Disease. *Journal of Alzheimer's Disease*, 12(3), 245–253.

Ullrich, S., Munch, A., Neumann, S., Kremmer, E., Tatzelt, J. & Lichtenthaler, S.F. (**2010**) The Novel Membrane Protein TMEM59 Modulates Complex Glycosylation, Cell Surface Expression, and Secretion of the Amyloid Precursor Protein. *The Journal of Biological Chemistry*, 285(27), 20664–20674.

Valinluck, V. & Sowers, L.C. (**2007**) Endogenous Cytosine Damage Products Alter the Site Selectivity of Human DNA Maintenance Methyltransferase DNMT1. *Cancer Research*, 67(3), 946–950.

Valinluck, V., Tsai, H.-H., Rogstad, D.K., Burdzy, A., Bird, A. & Sowers, L.C. (**2004**) Oxidative damage to methyl-CpG sequences inhibits the binding of the methyl-CpG binding domain (MBD) of methyl-CpG binding protein 2 (MeCP2). *Nucl. Acids Res.*, 32(14), 4100–4108.

Verghese, P.B., Castellano, J.M. & Holtzman, D.M. (**2011**) Apolipoprotein E in Alzheimer's disease and other neurological disorders. *The Lancet Neurology*, 10(3), 241–252.

Varley, K.E., Gertz, J., Bowling, K.M., Parker, S.L., Reddy, T.E., Pauli-Behn, F., Cross, M.K., Williams, B.A., Stamatoyannopoulos, J.A., Crawford, G.E., Absher, D.M., Wold, B.J. & Myers, R.M. (**2013**) Dynamic DNA methylation across diverse human cell lines and tissues. *Genome Research*.

Vázquez-Bourgon, J., Arranz, M.J., Mata, I., Pelayo-Terán, J.M., Pérez-Iglesias, R., Medina-González, L., Carrasco-Marín, E., Vázquez-Barquero, J.L. & Crespo-Facorro, B. (**2010**) Serotonin transporter polymorphisms and early response to antipsychotic treatment in first episode of psychosis. *Psychiatry Research*, 175(3), 189–194.

Verkaik, R., Nuyen, J., Schellevis, F. & Francke, A. (**2007**) The relationship between severity of Alzheimer's disease and prevalence of comorbid depressive symptoms and depression: a systematic review. *International Journal of Geriatric Psychiatry*, 22(11), 1063–1086.

Vertes, R.P., Fortin, W.J. & Crane, A.M. (**1999**) Projections of the median raphe nucleus in the rat. *The Journal of Comparative Neurology*, 407(4), 555–582.

Vostrov, A.A., Taheny, M.J., Izkhakov, N. & Quitschke, W.W. (**2010**) A nuclear factor-binding domain in the 5'-untranslated region of the amyloid precursor protein promoter: Implications for the regulation of gene expression. *BMC Research Notes*, 3(1), 1–6.

Waldemar, G., Dubois, B., Emre, M., Georges, J., McKeith, I.G., Rossor, M., Scheltens, P., Tariska, P. & Winblad, B. (**2007**) Recommendations for the diagnosis and management of Alzheimer's disease and other disorders associated with dementia: EFNS guideline. *European Journal of Neurology*, 14(1), e1–e26.

Wallace, K., Grau, M.V., Levine, A.J., Shen, L., Hamdan, R., Chen, X., Gui, J., Haile, R.W., Barry, E.L., Ahnen, D., McKeown-Eyssen, G., Baron, J.A. & Issa, J.P.J. (**2010**) Association between Folate Levels and CpG Island Hypermethylation in Normal Colorectal Mucosa. *Cancer Prevention Research*, 3(12), 1552–1564.

Wang, D., Szyf, M., Benkelfat, C., Provençal, N., Turecki, G., Caramaschi, D., Côté, S.M., Vitaro, F., Tremblay, R.E. & Booij, L. (**2012**) Peripheral SLC6A4 DNA Methylation Is Associated with In Vivo Measures of Human Brain Serotonin Synthesis and Childhood Physical Aggression. *PLoS ONE*, 7(6).

Wang, S.-C., Oelze, B. & Schumacher, A. (**2008**) Age-Specific Epigenetic Drift in Late-Onset Alzheimer's Disease. *PLoS ONE*, 3(7), e2698.

Weaver, I.C.G., Cervoni, N., Champagne, F.A., D'Alessio, A.C., Sharma, S., Seckl, J.R., Dymov, S., Szyf, M. & Meaney, M.J. (**2004**) Epigenetic programming by maternal behavior. *Nat Neurosci*, 7(8), 847–854.

Weaver, I.C.G., Champagne, F.A., Brown, S.E., Dymov, S., Sharma, S., Meaney, M.J. & Szyf, M. (**2005**) Reversal of Maternal Programming of Stress Responses in Adult Offspring through Methyl Supplementation: Altering Epigenetic Marking Later in Life. *J. Neurosci.*, 25(47), 11045–11054.

Weitzman, S.A., Turk, P.W., Milkowski, D.H. & Kozlowski, K. (**1994**) Free radical adducts induce alterations in DNA cytosine methylation. *Proceedings of the National Academy of Sciences of the United States of America*, 91(4), 1261–1264.

West, R., Lee, J. & Maroun, L. (**1995**) Hypomethylation of the amyloid precursor protein gene in the brain of an alzheimer's disease patient. *Journal of Molecular Neuroscience*, 6(2), 141–146.

Weuve, J., Korrick, S.A., Weisskopf, M.A., Ryan, L.M., Schwartz, J., Nie, H., Grodstein, F. & Hu, H. (**2009**) Cumulative Exposure to Lead in Relation to Cognitive Function in Older Women. *Environmental Health Perspectives*, 117(4), 574–580.

Whitwell, J.L., Przybelski, S.A., Weigand, S.D., Knopman, D.S., Boeve, B.F., Petersen, R.C. & Jack, C.R. (**2007**) 3D maps from multiple MRI illustrate changing atrophy patterns as subjects progress from mild cognitive impairment to Alzheimer's disease. *Brain*, 130(7), 1777–1786.

Willeit, M., Stastny, J., Pirker, W., Praschak-Rieder, N., Neumeister, A., Asenbaum, S., Tauscher, J., Fuchs, K., Sieghart, W., Hornik, K., Aschauer, H.N., Brücke, T. & Kasper, S. (**2001**) No evidence for in vivo regulation of midbrain serotonin transporter availability by serotonin transporter promoter gene polymorphism. *Biological Psychiatry*, 50(1), 8–12.

Wirak, D.O., Bayney, R., Kundel, C.A., Lee, A., Scangos, G.A., Trapp, B.D. & Unterbeck, A.J. (**1991**) Regulatory region of human amyloid precursor protein (APP) gene promotes neuron-specific gene expression in the CNS of transgenic mice. *The EMBO Journal*, 10(2), 289–296.

Wisniewski, K.E., Wisniewski, H.M. & Wen, G.Y. (**1985**) Occurrence of neuropathological changes and dementia of Alzheimer's disease in Down's syndrome. *Annals of Neurology*, 17(3), 278–282.

Wolf, H., Julin, P., Gertz, H.-J., Winblad, Bengt & Wahlund, L.-O. (**2004**) Intracranial volume in mild cognitive impairment, Alzheimer's disease and vascular dementia: evidence for brain reserve? *International Journal of Geriatric Psychiatry*, 19(10), 995–1007.

Wong, C.C.Y., Caspi, A., Williams, B., Craig, I.W., Houts, R., Ambler, A., Moffitt, T.E. & Mill, J. (**2010**) A longitudinal study of epigenetic variation in twins. *Epigenetics*, 5(6), 516–526.

Wright, R.O., Tsaih, S.W., Schwartz, J., Spiro, A., McDonald, K., Weiss, S.T. & Hu, H. (**2003**) Lead Exposure Biomarkers and Mini-Mental Status Exam Scores in Older Men. *Epidemiology*, 14(6), 713–718.

Wu, H., D'Alessio, A.C., Ito, S., Wang, Z., Cui, K., Zhao, K., Sun, Y.E. & Zhang, Yi (**2011**) Genome-wide analysis of 5-hydroxymethylcytosine distribution reveals its dual function in transcriptional regulation in mouse embryonic stem cells. *Genes & Development*, 25(7), 679–684.

Wu, H.-C., Delgado-Cruzata, L., Flom, J.D., Kappil, M., Ferris, J.S., Liao, Y., Santella, R.M. & Terry, M.B. (**2011**) Global methylation profiles in DNA from different blood cell types. *Epigenetics: Official Journal of the DNA Methylation Society*, 6(1).

Wu, J., Basha, M.R., Brock, B., Cox, D.P., Cardozo-Pelaez, F., McPherson, C.A., Harry, J., Rice, D.C., Maloney, B., Chen, D., Lahiri, D.K. & Zawia, N.H. (**2008**) Alzheimer's Disease (AD)-Like Pathology in Aged Monkeys after Infantile Exposure to Environmental Metal Lead (Pb): Evidence for a Developmental Origin and Environmental Link for AD. *J. Neurosci.*, 28(1), 3–9.

Wyss-Coray, T. & Mucke, L. (**2002**) Inflammation in neurodegenerative disease--a double-edged sword. *Neuron*, 35(3), 419–432.

Yamada, T., Sasaki, H., Dohura, K., Goto, I. & Sakaki, Y. (**1989**) Structure and expression of the alternatively-spliced forms of mRNA for the mouse homolog of Alzheimer's disease amyloid beta protein precursor. *Biochemical and Biophysical Research Communications*, 158(3), 906–912.

Yen, P.H., Patel, P., Chinault, A.C., Mohandas, T. & Shapiro, L.J. (**1984**) Differential methylation of hypoxanthine phosphoribosyltransferase genes on active and inactive human X chromosomes. *Proceedings of the National Academy of Sciences*, 81(6), 1759–1763.

Yu, M., Hon, G.C., Szulwach, K.E., Song, C.-X., Jin, P., Ren, B. & He, C. (**2012**) Tet-assisted bisulfite sequencing of 5-hydroxymethylcytosine. *Nature Protocols*, 7(12), 2159–2170.

Zawia, N.H., Lahiri, D.K. & Cardozo-Pelaez, F. (**2009**) Epigenetics, oxidative stress, and Alzheimer disease. *Free Radical Biology and Medicine*, 46(9), 1241–1249.

Zhou, H., Hu, Hu & Lai, M. (**2010**) Non-coding RNAs and their epigenetic regulatory mechanisms. *Biology of the Cell*, 102(12), 645–655.

Zhu, C.-H., Huang, Yuanhui, Oberley, L.W. & Domann, F.E. (**2001**) A Family of AP-2 Proteins Down-regulate Manganese Superoxide Dismutase Expression. *Journal of Biological Chemistry*, 276(17), 14407–14413.

Zhu, X., Lee, H., Casadesus, Gemma, Avila, J., Drew, K., Perry, George & Smith, M. (**2005**) Oxidative imbalance in alzheimers disease. *Molecular Neurobiology*, 31(1-3), 205–217.

Ziller, M.J., Müller, F., Liao, J., Zhang, Yingying, Gu, H., Bock, C., Boyle, P., Epstein, C.B., Bernstein, B.E., Lengauer, T., Gnirke, A. & Meissner, A. (**2011**) Genomic Distribution and Inter-Sample Variation of Non-CpG Methylation across Human Cell Types. , 7(12).

Zukin, S. (**2009**) Epigenetics. *Alzheimer's and Dementia*, 5(4, Supplement 1), P146–P147.

APPENDIX A

This appendix lists the samples used in each of the results chapters. The following table summarizes and sample demographics for the samples used in Chapter 3 and briefly described in section 3.3:

Disease Status	Sex	Age	ApoE Genotype	Braak Stage	Sample ID
AD	Female	96	3,4	4	8
AD	Female	92	3,4	6	9
AD	Female	89	3,3	6	11
AD	Female	80	3,4	5	13
AD	Female	91	3,4	3	16
AD	Female	84	3,4	6	17
AD	Female	82	4,4	6	18
AD	Female	81	3,4	-	19
AD	Female	69	3,3	6	20
AD	Female	90	3,3	-	21
AD	Female	91	4,4	6	22
AD	Female	96	4,4	5	23
AD	Female	85	4,4	5	24
AD	Female	97	3,3	5	25
AD	Female	87	3,3	3	26
AD	Female	85	3,3	6	27
AD	Female	87	2,4	6	29
AD	Female	91	2,4	5	31
AD	Female	92	3,4	6	33
AD	Female	94	4,4	5	35
AD	Female	88	3,4	6	38
AD	Female	90	3,4	6	39
AD	Female	80	4,4	6	40
AD	Female	93	3,4	6	41
AD	Female	73	3,3	0	54
AD	Female	98	3,3	3	57
AD	Female	98	3,4	6	59
AD	Male	89	3,3	-	3
AD	Male	82	3,3	6	10
AD	Male	91	3,3	6	12
AD	Male	89	2,3	-	14
AD	Male	84	4,4	6	15
AD	Male	80	3,4	6	30
AD	Male	81	4,4	6	32

AD	Male	71	2,3	6	34
AD	Male	90	3,4	6	36
AD	Male	86	2,4	5	37
AD	Male	80	3,4	3	56
AD	Male	72	3,3	6	58
Ctrl	Female	82	3,3	0	1
Ctrl	Female	82	2,3	2	2
Ctrl	Female	81	3,3	1	4
Ctrl	Female	92	3,3	2	5
Ctrl	Female	80	3,3	0	28
Ctrl	Female	55	3,3	1	42
Ctrl	Female	88	3,4	5	45
Ctrl	Female	87	3,3	0	48
Ctrl	Female	68	-,-	0	49
Ctrl	Female	80	3,4	5	51
Ctrl	Female	90	3,3	2	61
Ctrl	Female	88	3,3	6	64
Ctrl	Female	90	2,3	3	69
Ctrl	Female	89	3,3	4	77
Ctrl	Female	92	3,4	1	94
Ctrl	Male	78	2,3	1	6
Ctrl	Male	86	3,4	3	7
Ctrl	Male	79	2,3	2	43
Ctrl	Male	80	3,3	1	44
Ctrl	Male	86	3,3	0	47
Ctrl	Male	59	4,4	0	50
Ctrl	Male	40	3,4	0	52
Ctrl	Male	81	2,4	3	60
Ctrl	Male	81	3,3	1	62
Ctrl	Male	66	3,3	1	63
Ctrl	Male	86	3,3	3	65
Ctrl	Male	86	3,3	3	90

The following table summarizes and sample demographics for the samples used in Chapter 4 and briefly described in section 4.3:

Disease Status	Sex	Age	ApoE Genotype	Braak Stage	Sample ID
AD	Female	80	3,4	5	13
AD	Female	88	3,4	6	38
AD	Female	80	4,4	6	40
AD	Female	96	3,4	4	8
AD	Female	92	3,4	6	9
AD	Female	89	3,3	6	11

AD	Female	91	3,4	3	16
AD	Female	84	4,4	6	17
AD	Female	81	3,4	-	19
AD	Female	69	3,3	6	20
AD	Female	90	3,3	-	21
AD	Female	96	4,4	5	23
AD	Female	85	4,4	5	24
AD	Female	97	3,3	5	25
AD	Female	87	3,3	3	26
AD	Female	87	2,4	6	29
AD	Female	91	2,4	5	31
AD	Female	92	3,4	6	33
AD	Female	94	4,4	5	35
AD	Female	93	3,4	6	41
AD	Female	98	3,3	3	57
AD	Female	98	3,4	6	59
AD	Female	82	4,4	6	18
AD	Female	91	4,4	6	22
AD	Female	85	3,3	6	27
AD	Female	90	3,4	6	39
AD	Female	73	3,3	0	54
AD	Male	82	3,3	6	10
AD	Male	84	4,4	6	15
AD	Male	90	3,4	6	36
AD	Male	86	2,4	5	37
AD	Male	91	3,3	6	12
AD	Male	89	2,3	-	14
AD	Male	80	3,4	6	30
AD	Male	81	4,4	6	32
AD	Male	71	2,3	6	34
AD	Male	72	3,3	6	58
Ctrl	Female	82	3,3	0	1
Ctrl	Female	82	2,3	2	2
Ctrl	Female	81	3,3	1	4
Ctrl	Female	92	3,3	2	5
Ctrl	Female	80	3,3	0	28
Ctrl	Female	55	3,3	1	42
Ctrl	Female	88	3,4	5	45
Ctrl	Female	87	3,3	0	48
Ctrl	Female	68	-,-	0	49
Ctrl	Female	80	3,4	5	51
Ctrl	Female	55	3,4	0	53
Ctrl	Female	90	3,3	2	61
Ctrl	Male	89	3,3	-	3
Ctrl	Male	78	2,3	-	6
Ctrl	Male	86	3,4	3	7

Ctrl	Male	79	2,3	2	43
Ctrl	Male	80	3,3	1	44
Ctrl	Male	78	2,3	1	46
Ctrl	Male	86	3,3	0	47
Ctrl	Male	59	4,4	0	50
Ctrl	Male	40	3,4	0	52
Ctrl	Male	80	3,4	3	56
Ctrl	Male	81	2,4	3	60
Ctrl	Male	81	3,3	1	62

The following table summarizes and sample demographics for the samples used in Chapter 5 and briefly described in section 5.3:

Disease Status	Sex	Age	ApoE Genotype	Braak Stage	Sample ID	Passed QC
AD	Female	82	2,3	2	2	Yes
AD	Female	96	3,4	4	8	Yes
AD	Female	90	3,4	-	9	Yes
AD	Female	89	3,3	6	11	Yes
AD	Female	80	3,4	5	13	No
AD	Female	91	3,4	6	16	Yes
AD	Female	84	3,4	6	17	Yes
AD	Female	82	4,4	6	18	No
AD	Female	81	3,4	-	19	Yes
AD	Female	69	3,3	6	20	Yes
AD	Female	90	3,3	-	21	No
AD	Female	91	4,4	6	22	Yes
AD	Female	96	4,4	5	23	Yes
AD	Female	85	4,4	5	24	Yes
AD	Female	97	3,3	5	25	Yes
AD	Female	85	3,3	6	27	Yes
AD	Female	87	2,4	6	29	Yes
AD	Female	91	2,4	5	31	Yes
AD	Female	92	3,4	6	33	Yes
AD	Female	94	4,4	5	35	Yes
AD	Female	88	3,4	6	38	Yes
AD	Female	90	3,4	6	39	Yes
AD	Female	80	4,4	6	40	Yes
AD	Female	93	3,4	0	41	Yes
AD	Female	98	3,3	3	57	Yes
AD	Female	98	3,4	6	59	Yes
AD	Female	87	3,3	3	66	Yes
AD	Female	79	3,4	6	75	No
AD	Female	85	3,4	6	76	Yes

AD	Female	89	3,3	4	77	No
AD	Female	85	3,4	6	79	Yes
AD	Female	84	3,3	4	80	Yes
AD	Female	86	3,4	6	81	Yes
AD	Female	103	3,3	5	82	Yes
AD	Female	81	3,4	6	85	Yes
AD	Female	83	3,3	3	89	No
AD	Female	84	3,4	6	91	Yes
AD	Female	92	3,4	5	92	Yes
AD	Female	80	4,4	6	96	Yes
AD	Female	98	3,3	6	97	Yes
AD	Female	83	3,4	6	106	Yes
AD	Male	89	3,3	-	3	Yes
AD	Male	82	3,3	-	10	Yes
AD	Male	91	3,3	6	12	Yes
AD	Male	89	2,3	-	14	Yes
AD	Male	84	4,4	6	15	Yes
AD	Male	80	3,4	6	30	Yes
AD	Male	81	4,4	6	32	No
AD	Male	71	2,3	6	34	Yes
AD	Male	90	3,4	6	36	Yes
AD	Male	86	2,4	5	37	Yes
AD	Male	80	3,4	3	56	Yes
AD	Male	72	3,3	6	58	Yes
AD	Male	81	2,4	3	60	Yes
AD	Male	86	3,3	3	65	Yes
AD	Male	82	3,4	4	67	Yes
AD	Male	88	3,4	5	71	Yes
AD	Male	80	3,3	6	72	Yes
AD	Male	74	3,4	6	83	Yes
AD	Male	88	3,3	6	84	Yes
AD	Male	77	3,4	5	86	Yes
AD	Male	86	3,4	6	87	Yes
AD	Male	75	4,4	6	88	No
AD	Male	97	3,4	5	95	Yes
AD	Male	88	2,4	6	100	Yes
AD	Male	105	3,4	5	103	Yes
Ctrl	Female	82	3,3	0	1	No
Ctrl	Female	81	3,3	1	4	No
Ctrl	Female	92	3,3	2	5	Yes
Ctrl	Female	80	3,3	0	28	Yes
Ctrl	Female	55	3,3	0	42	Yes
Ctrl	Female	88	3,4	5	45	Yes
Ctrl	Female	87	3,3	0	48	Yes
Ctrl	Female	68	-,-	0	49	Yes
Ctrl	Female	80	3,4	0	51	Yes

Ctrl	Female	55	3,4	0	53	Yes
Ctrl	Female	73	3,3	0	54	Yes
Ctrl	Female	90	3,3	2	61	No
Ctrl	Female	88	3,3	6	64	Yes
Ctrl	Female	90	2,3	3	69	Yes
Ctrl	Female	85	3,3	1	99	No
Ctrl	Female	94	-,-	4	102	Yes
Ctrl	Female	95	-,-	5	104	Yes
Ctrl	Male	86	3,4	3	7	No
Ctrl	Male	79	2,3	2	43	No
Ctrl	Male	80	3,3	0	44	No
Ctrl	Male	78	2,3	1	46	Yes
Ctrl	Male	86	3,3	0	47	Yes
Ctrl	Male	59	4,4	0	50	Yes
Ctrl	Male	40	3,4	0	52	Yes
Ctrl	Male	81	3,3	1	62	Yes
Ctrl	Male	66	3,3	0	63	Yes
Ctrl	Male	86	3,3	3	90	Yes
Ctrl	Male	97	-,-	2	101	Yes

APPENDIX B

This appendix lists the target-specific oligo primers used for the APP bisulphite PCR sequencing described in Chapter 5 and covered by the Chapter 2 materials and methods section 2.18. All primers are tagged with the CS sequences as follows:

CS Tag	Sequence (5'-3')
Forward primer tag	ACACTGACGACATGGTTCTACA–[Fwd]
Reverse primer tag	TACGGTAGCAGAGACTTGGTCT–[Rev]

Access Array set	Name	Forward Sequence (5'-3')	Reverse Sequence (5'-3')
1	APP_3	GGAAAGAGAAATTGTTAAATTTAAGAAT	AATCAACACCATCCTAACTAACAC
1	APP_11	TAAACAATAAAAAAAAAAATCTAAAACCAAA	GGATAAAAAYGTATTTTAGTAGTAGTTTTT
1	APP_76	AAACATATTTTTAATTAAAAAAAACAATAA	ATTATAGAGTTTGTGGAAGAGGTG
1	APP_12	GGTTTTAGATTTTTTTTTTTATTGTTTA	ACACCCRAAAAAAACCCCTAAC
1	APP_91	ATAATATAATTCTCCAAAACCAAA	TGTTGTTTTATTTTATTTATTTAGTTTTT
1	APP_15	GTTTGGTATTGTTTTTGTTGGT	TTCTCTACATTAAAAACTTAAATTAAAATCTTA
1	APP_111	TTTAGTATGTATTAATATGAATTAATTAGAATG	AAACTCATAAAAAAAAATCTATCTAACC
1	APP_136	AGAATAATGTGGGAAGAAATAAATT	ACAAAACCCATTAATAATATAATATAAAAAC
1	APP_141	TTTTGTTGGTTTTTAATTATTAAGAATTT	ATCTATAATCATCCTTCAAAAAAAA
1	APP_146	TTAAAATGGAGTTGTTTTGGTTTAG	CCATTAAAAAAATACTATCAATATACCTCA
1	APP_14	GGATTAGTTGATTYGTTTGGTTTTGAGTTT	ACCAACAAAAACAATACCAAAC
1	APP_20	GGGATAGTTTAGGAAGTTAGGAGTT	CATACACAAAATTTACCTATAAAAAAC
1	APP_110	TGTTTAGAGGGAAAATATTTTTGTATT	ACAATTCTCCTACCTCAACCTC
1	APP_121	TGTTTAGTTGAGTTTGTTGATGT	TTTAATTTCTAAAATCACTTCAAATT
1	APP_124	GTTGTTGTTGAAGAGGTTAATAAAAAT	TAAACTTATATAACACAAACCCCAAA
1	APP_145	TTAATGGTTTGGGGATTAATTAATA	CATATTAACCACAACAAAATCTCC
1	APP_1	TAGATAAGTGGTTGTTGGTTTTTTT	TCCTAATAACTTTTTACTTTTCAACC
1	APP_2	GGATTTTTGAGTAGTGAAGGAATTA	TCTTAAATTTAACAATTTCTCTTTCC
1	APP_4	TGATTTTAGGTGATTTGTTTATTTG	AAACCCAATTATTCACTTTCTACAC
1	APP_6	TAGGATAGATATAATGAAGAATAAGGGTAG	AAACAAAAATATAACATCCATCTAAATTAA
1	APP_10	GGATAAAAAYGTATTTTAGTAGTAGTTTTT	ATACAAAATCAAAAAAAAATAAATCCTAAAAC
1	APP_13	GTTAGGGGTTTTTTTYGGGTGT	AAACTCAAAACCAAACRAATCAACTAATCC
1	APP_19	TTTTTAAAGTTAGTTTTTYGTTTTTGAA	AACTCCTAACTTCCTAAACTATCCC
1	APP_23	GTTTGGGAGGGGGTAGAGA	CCAAAAACTTTACAACCCAACACA
1	APP_28	TTTGAAGATGTTTAGAAGTTGTTTGT	TTTTTCTTTTAAAAAATTACTAAAAAAC
1	APP_45	GTTTATTATAAATTTAAGGTTATTTGGA	TTAATACAAAAATCTAACAAAAATACA
1	APP_75	ATAGTGAAGATAAAGTAGTAGAAGTAGTAGAGG	AATAATAATAATAACAATACTAATAATTCTCTC
1	APP_113	GTTATGGTAATGTTGTGGGTAAGTT	TCTCTAAAAAATAAATCACATACAAAAAA
1	APP_117	TTTGGTTAATATGATTAGTGAATTAAGGAT	ATCAACCCCAAAAAAATACCAC
1	APP_140	ATAATTAAAAACCAACAAAACACCC	ATTAGGTATTGAGATTTTAAGTTTTTT

1	APP_6.1	ACCAACCCATTTACTTCTTCTTAAAAAC	TTAATTTAGATGGATGTTATATTTTTGTTTAAT
1	APP_**18.1**	TCCCCCCTATCTAAATATCTCTC	AGGGTAGGGGGYGATTTTGAG
1	APP_5	TGAATAATTGGGTTTGTTTTGTTAT	AAATTACTTTTACCCCTAAATATCCTTTA
1	APP_8	TTTGTTTTTAAGAAGAAGTAAATGG	ATTAAAACAACAAAACCTCATTTTT
1	APP_9	AAAATGAGGTTTTGTTGTTTTAATAAG	AAAAACTACTACTAAAATACRTTTTTATCC
1	APP_16	TAAGATTTTAATTTAAGTTTTTAATGTAGAGAA	ACTCCCTAARACAAAACTAAACC
1	APP_17	GGTTTAGTTTTGTYTTAGGGAGT	ACTCAAAATCRCCCCCTACC
1	APP_18.2	ATTCAAAAACRAAAAACTAACTTTAAAAAC	GAGAGATATTTAGATAGGGGGGA
1	APP_21	TTTTTTATAGGTAAATTTTGTGTATGTT	ACCTTCTACCTCTCAAACAAAACT
1	APP_22	AGTTTTGTTTGAGAGGTAGAAGGT	TCTCTACCCCCTCCCAAAC
1	APP_24	TGTGTTGGGTTGTAAAGTTTTTGG	AAACTAAACCAAAATCCCCAA
1	APP_32	TTGATTGTTTTTTATAGAGTTTAGAAA	AAATTTATTAAACTCACACCTCTATC
1	APP_34	ATTTCCTCAATCAAAACCAATATTA	TAGTAAAGTTTAGGTTTGATGGTGATT
1	APP_36	CTACTAAATTTAATACACTTCTCACTAATAATT	TTTTTGTATATTAGAGTTTTTTATAGATATTTG
1	APP_81	TTGTTGGAGGGATTAAATTTAAGTG	CCTAAATAATCTTAAATAAACTTTAAAACAC
1	APP_135	AATTTATTTCTTCCCACATTATTCT	TTTTTTGAGTAGATGTAGAATTAGATTTT
1	APP_137	TTAATTAAATCACAAACCACAAAAATAATATAC	TTTGTGTATTGTAAAGAATTTAGTTGTATTAAA
1	APP_25	GGGTTTTGTTTTGTGTATTTAAGGA	ACCCTCCCCAACATCCTATAC
2	APP_26	GGTTTAGTTTTTTTGGGGGTATA	CCACACAACTTTTCTTTACTACACTAC
2	APP_27	TGTGGAGGTTGTAGTTTAGTGAAA	ACAAACAACTTCTAAACATCTTCAAA
2	APP_29	GTTTTTTAGTAATTTTTTAAAAGAAAAA	TCCTAAACCACTTATAAATAAAAATCAA
2	APP_30	TTTTTTTTGATTTTTATTTATAAGTGGTT	ATCAAAATTCTCTTAATTAAATACATTTTT
2	APP_31	GGGTGGAGGTGAGATTTATATTATGT	ACCTATTTTCTAAACTCTATAAAAAACAATCAA
2	APP_33	TTTTTTTTGTTGAGGGTTTAATTTA	AATCACCATCAAACCTAAACTTTACTA
2	APP_35	TAATATTGGTTTTGATTGAGGAAAT	ACAACAAAACTTAACATAAAAACAAC
2	APP_37	AAAGTTTTTAGAATGTAAGGTTTAGTT	ATTCCACTTTATTTTATTATCATATACATATTT
2	APP_38	GTGATGGAATAAAGTAATGATTTTTG	TTACTTTCAAAAAACAAACACAACA
2	APP_39	TGAGGTTTTTAGAAAAGAGTTTGATAT	TCTATTTCAATAACACCAACAAACA
2	APP_40	AAGAAAATTTTATAAAAGGAAATTGATT	AATCAACAACATATCAACACTAACC
2	APP_41	GGTTAGTGTTGATATGTTGTTGATTTA	AACAACTATAAAATACTCCCTAAAA
2	APP_42	TTTTGAAAAGAGTTAAAATTGTTATTAAAT	CCAAATTATCTCTAAAAAAAACAAACTA
2	APP_43	TAGTTTGTTTTTTTTAGAGATAATTTGG	AACTATCCCTACAAAAATACCTACTATT
2	APP_44	AATAGTAGGTATTTTTGTAGGGATAGTT	CTTTAAAAACTAATCATATTATCTCAACTTATA
2	APP_46	TTTTAGTATTTTTGGTTAGTTAATTTGGT	CATTTTCCAAACAATAAACATATTC
2	APP_47	TGTTTGGAAAATGTTTATTTGAATATA	CAAAACAACAAATACTAATAATTATCAAAC
2	APP_48	TTATTAGTATTTGTTGTTTTGTTTAGTTATTAA	AAAACTAATTATCAAAAATTTCAATATAACC
2	APP_49	TTAGGAAATTATGGATGTGATAGAAATA	CCTAAATTCAAACTATTCTCCTACCT
2	APP_50	AAAATATTGAGATTATGTAAGAATTTGATT	CCATCTATTTTCTACATTTAAACCAC
2	APP_51	GATGGTTTTTATTAGTTAGAAATTTTAAAATTA	AAAAAAAATAAACAAAAAACTCTCTATTTTACT
2	APP_52	TTTTTTTTGTTTTTAAGATTTGTAGTATTT	ACTTTCCCATAAAAAAACAATATTC
2	APP_53	ATGTTAGTGGTAAGTTATTTAGTTTATAAGAAT	AAATTAAAATTACTTTACAAAAATCTAAAACA
2	APP_54	TTTAATTTAAAGGAGTGTTGAAGAT	TATCCATAACTCCAAAACAAAAACT
2	APP_55	TTAAGTTTTTGTTTTGGAGTTATGG	CACTTCCCATTCTAAACATTCATATA
2	APP_56	ATATGAATGTTTAGAATGGGAAGTG	TTTTAAAAACAAATACATATAATCCAAC
2	APP_57	TTTAATTTTGAGTATGAAAGTGTGG	TTTCCAAATAATAACCTTATATAAAAACAA
2	APP_58	AAGGTTATTATTTGGAAATATAGTATGTTG	CTAAAAATACAATCTCTAACTTAATTAAAAAAA

2	APP_59	GTTTAGGATATGTTTTGTTTTGGTT	CCTAAATTAATAATTACTCCAATTTATTTAAA
2	APP_60	TATGTTTTTTAGGGTGGATTTTT	ACCACTAAAAATTTAATATTCCCATAAC
2	APP_61	TGTGGTAGAAGTTAATTAATTAGTGATTATT	TTCTTCCCTCAAAACCAAAC
2	APP_62	TAGTTTGAAGATGTTTGAAGTTTGA	ACACACCAAACAAAAAATAAATACAT
2	APP_63	TTTTTGTATTTAAATAAATGTATTTATTTTTTG	TTCCTTTACTACTTTTCTACTTAACTAAATTTT
2	APP_66	AAGTAGAGTTATTTTGGGAAGGTTT	AATTTTCAAACTCAAACCATACTTC
2	APP_67	AATTTGTTTTGAAGTTTATTTTTTTT	CAAACATCCATCCTCTCCTAATATAA
2	APP_68	TATATTAGGAGAGGATGGATGTTTG	AACTATTACCTCAAAATACCCCTAAA
2	APP_69	TTTAGATTTTAGTTTTTAGAGTGATTTTAG	AATTCAACTAACCCTATTTTTAACTTTT
2	APP_70	TTTTTTTTATATTTTGTAGATATGTAGTGAGA	ACAAAATCCACATTATCACTTTCTT
2	APP_71	ATTGGTTGAAGAAAGTGATAATGT	CATCTCTAAAACTAAACACAAAAACC
2	APP_72	GTTGTTTTTTTGTATTGTATTTTGAGA	CCACTCTTACTTAACTCATCCACTAA
2	APP_73	GTTGTTATAATGTTGGGTTGGTAGT	AACCAATTCCCTCAAAAAAAA
2	APP_74	GGGATTTATATTTAGATTTGTTGAATT	TTTATCTTCACTATAAAAACAAACACAA
2	APP_77	AATAGATTGATTTTTTTGGTTTTTG	AAACACATCCATAACCATACTTTATC
2	APP_78	TGTTTTTTTATGTTAAATGTGGTTTT	AACACACTTCCCTTCAATCAC
2	APP_79	TGATTGAAGGGAAGTGTGTTT	AATCTATTACAAAAAACAAAATCAAT
2	APP_80	TGATTGGATGATTGTAGTTTTTG	TCACTATAACAAAAAACCTTTATTCC
2	APP_82	AAGTTTATTTAAGATTATTTAGGAATTTTTT	TATTCAAAAAACACAAAACCATACA
2	APP_83	AGTTTTGAGATTTGAATAGGATGTATTA	CATTTAAATTTAAATCCCATTTCTTT
3	APP_84	ATTTTAATGTTATAATTGTAAAATGGATTA	CAAAACAAAAAATTTAATAAAACAAC
3	APP_85	AATTTTTTGTTTTGAGAAATAATTGAAAA	CATCAAAAATACTAACTACTATTATAAAAACT
3	APP_86	TTTTTTTTATTTTATAGTTTTTATAATAGTAGT	CTTAACCTCAAACCTCTCTTTAACTT
3	APP_87	TTAGAAAGTTAAAGAGAGGTTTGAG	TTAAAAATCTTAATAAAACCAACAAAC
3	APP_88	AAAATGATAAAAATAGAGGAAAAATATAATTAT	CCTTCTTATCAACTTTAAACAAATTCT
3	APP_89	AAGAAGGTAGTTATTTAGGTAAAATTTGA	AAATTCATAAAACTATAAACAATCACAC
3	APP_90	TTTATGAAATATTGATTTTTAATGAGAAA	AAACAACAATTTCAAATCTTCTAAAA
3	APP_92	GTTTGGTTTTGGAGAATTATATTAT	TACTCAAAACATTTCTCCTCATTCT
3	APP_93	TTTGATGTTTAATGTTTTGTAGAAATG	CTCTACCTCCCAAATTCAAATAAT
3	APP_94	GGTAGGATTATGTTTAGGGTTTTTAA	TTATCCCACTCCACAACTTTATACT
3	APP_95	TTGTTTTTGAGGGATAGAGATAATTAT	CATTAATTTAACCACCTAAACACATC
3	APP_96	GGTTTTTTTGAATTATTTGATGTTT	ACCTAAAAAACCCCTTCACTATTTA
3	APP_97	AATGTTGAGTAAGGAGGATAGTTATTATATAT	AAATCACCTTTTTACAAATCTACCTT
3	APP_98	AATGTTTAGAAGGTAGATTTGTAAAAAG	ATACAAAACACTACTTCTTCCATTAAC
3	APP_99	GTGTTTTGTATAAATTTGAAAAAGAGTT	AAATACTTTAAAATATACTATCTATCCTTCTA
3	APP_100	AGAAGGATAGATAGTATATTTTAAAGTATTT	TATCACCAACCCAAAAAACAA
3	APP_101	TTGTTTTTTGGGTTGGTGATA	CCAAAAATTAAAAAACTCTTACCATAA
3	APP_102	AGTGTTTGGTAAGAAAAATAATTTG	AACCTACATCAAAAAATAACAACTC
3	APP_103	TGTTGATTTGGAGTTGTTATTTTTT	AAACAACTTACTTACCAACTTCATCCTA
3	APP_104	TGGTAAGTAAGTTGTTTTTTTGATGTT	CTAATTCCAACCTAACAATTTCATC
3	APP_105	AAATTGTTAGGTTGGAATTAGTTT	CCAAAACATCATAATTTAAAAACCATAA
3	APP_106	AAAATGAAAATTTTTAGATTATTGGAG	ACCTATATCACAACTAAATTTACTAAAATCA
3	APP_107	TTTTTTTGGTATTAGAGGTGGTG	TTAAATCATTTCCTCTCCAATAAAC
3	APP_108	AAGGGTATTGGAAGAGGAATAGAGT	CCTTTACTCAAAACTAAAATTAAACAAAAA
3	APP_112	AGATTTTTTTTATGAGTTTGAGAGA	CCACAACATTACCATAACAACTATAC
3	APP_114	TTGTGGATTGATTTTTAGAAAATTT	TCCTCAAATATTAATTTTCATACTCTTAC

3	APP_115	TTTTTATATTTGTAAAAGGGAGATGATA	AAATAACAACAATCACAACAACAAC
3	APP_116	GTTGTTGTTGTGATTGTTGTTATTT	ATCCTTAATTCACTAATCATATTAACCAAA
3	APP_118	TATTTTTTTGGGGTTGATTTTG	AAACATACAAAAAATAACTCTCTCTCTTAC
3	APP_119	GGTTTTTATTAGAAAATTAAGGAGTAG	AACATATACAATAAACAAATTACAAAAA
3	APP_120	TTTTTGTAATTTGTTTATTGTATATGTT	CATCAACAAACTCAACTAAACAC
3	APP_122	ATGAGTAAAATTAATTGGTTGTTTT	CCAATTTTTAATAATAAACTTCATATC
3	APP_123	GGATATGAAGTTTATTATTAAAAATTGGTA	TAAACAAACATTATATTTTTATCTTTTCCT
3	APP_125	TGGGGTTTGTGTTATATAAGTTTATTT	CCTTTAAAAAAATACTATAATTTTATAATATTC
3	APP_126	TGTATTTAAGAAATGAAATTTTTTTAATT	ATTACACCTTTATTTAAACCCACA
3	APP_127	TGGGTTTAAATAAAGGTGTAATTATT	AATTCCCACTTAAAAACATACAATC
3	APP_128	GTGGAGGTAGGTAAATTTGATTGTA	AAAAAACAACTATAACCCAAACATC
3	APP_129	AATTATTTGAATTTGGGAGGTAGA	ATATACAACTCCCTCATCCCTTAA
3	APP_130	TTTAAGGGATGAGGGAGTTGTA	TCAATATAATCAAATCCCCTATAATC
3	APP_131	GGGATTATAGGGGATTTGATTATAT	CCAACTACTCCAAAAACTAAAACA
3	APP_132	TTGGGTTTATTTATTTATTGTGGAG	CCCACCAATTACACAAATACTATACT
3	APP_133	TAGTATTTGTGTAATTGGTGGGTAA	AAAATAAAAATCAATCTTTTAAAAATAAA
3	APP_134	GTATTTTTTATTTTTTTATTGTTTTTTTAT	TACTCAAAAAACTTATAAATTAAATTTTC
3	APP_138	TTTTGTGGTTTGTGATTTAATTAAGT	AAACATCTAAAATACTTAAAAATATTTAACTTT
3	APP_139	AAAGTTAAATATTTTTAAGTATTTTAGATGTTT	AAAAAACTTAAAATCTCAATACCTAAT
3	APP_142	TTTTTTTTGAAGGATGATTATAGAT	TTTATCCAAACATACCTTCCTCATC
3	APP_143	GGTAGATTTAATTTTTTTTAATTAGTTTG	AATAATACTCCTCCAAAAATATATTTATTT
3	APP_144	AAATAAATATATTTTTGGAGGAGTATTATT	CTCAAACTTACAATTAAAACACAACA